Structural insights into the interaction of helicase and primase in Mycobacterium tuberculosis

Sharm, Dhakaram Pangeni; Vijayan, Ramachandran; Rehman, Syed Arif Abdul; Gourinath, Samudrala

Published in:
Biochemical Journal

DOI:
10.1042/BCJ20180673

Publication date:
2018

Document Version
Peer reviewed version

Structural insights into the interaction of helicase and primase in *Mycobacterium tuberculosis*

Dhakaram Pangeni Sharma[1], Ramachandran Vijayan[1], Syed Arif Abdul Rehman[2], Samudrala Gourinath[1]*

[1]School of Life Sciences, Jawaharlal Nehru University, Delhi-110067, India.

[2]MRC, Protein Phosphorylation & Ubiquitylation Unit, University of Dundee, Dundee, UK

*To whom correspondence should be addressed: Dr. Samudrala Gourinath, School of Life Sciences, Jawaharlal Nehru University, Delhi-110067, India.

Email: sgourinath@mail.jnu.ac.in, Tel. No: +91-11-26704513; Fax: +91 -11-26742916/2558

Keywords: DnaG primase, DnaB helicase, *Mycobacterium tuberculosis*, X-ray Crystallography, Molecular Dynamics, Protein-Protein Interactions.

Abstract:

The helicase-primase interaction is an essential event in DNA replication and is mediated by the highly variable C-terminal domain of primase (DnaG) and N-terminal domain of helicase (DnaB). To understand the functional conservation despite the low sequence homology of the DnaB-binding domains of DnaGs of eubacteria, we determined the crystal structure of the helicase-binding domain of DnaG from *M. tuberculosis* (*Mt*DnaG-CTD), and did so to a resolution of 1.58 Å. We observed the overall structure of *Mt*DnaG-CTD to consist of two subdomains, the N-terminal globular region (GR) and the C-terminal helical hairpin region (HHR), connected by a small loop. Despite differences in some of its helices, the globular region was found to have broadly similar arrangements across the species, whereas the helical hairpins showed different orientations. To gain insight into the crucial helicase-primase interaction in *M. tuberculosis*, a complex was modeled using the *Mt*DnaG-CTD and *Mt*DnaB-NTD crystal structures. Two nonconserved hydrophobic residues (Ile605 and Phe615) of *Mt*DnaG were identified as potential key residues interacting with *Mt*DnaB. Biosensor binding studies showed a significant decrease in the binding affinity of *Mt*DnaB-NTD with the Ile605Ala mutant of *Mt*DnaG-CTD compared to native *Mt*DnaG-CTD. The loop, connecting the two helices of the HHR, was concluded to be largely responsible for the stability of the DnaB-DnaG complex. Also, *Mt*DnaB-NTD showed micromolar affinity with DnaG-CTDs from *E. coli* and *H. pylori* and unstable binding with DnaG-CTD from *V.*

cholerae. The interacting domains of both DnaG and DnaB demonstrate the species-specific evolution of the replication initiation system.

Introduction:

DNA replication is a fundamental process in all domains of life and requires a large molecular machine coordinating the action of dozens of discrete factors to ensure accurate genome inheritance (1). The interaction between the bacterial replicative ring helicase (DnaB) and the primase (DnaG) is essential for the formation of the functional primosome at oriC to initiate the DNA replication (2, 3). DnaB helicase unwinds duplex DNA into single-stranded DNA (ssDNA), with this unwinding fueled by the hydrolysis of nucleoside triphosphate at the replication fork (4-6); and DnaG primase uses the newly formed ssDNA as a template for the de novo synthesis of RNA primers to initiate the DNA replication (7). The interaction between DnaB and DnaG mutually stimulates both of their activities (8). DnaG increases both the NTPase and helicase activities of DnaB (9) while DnaB increases and modulates the synthesis of RNA primers by DnaG (7, 10, 11). The DnaG primase consists of three domains with distinct activities; an N-terminal zinc-binding domain, a central RNA polymerase domain and a C-terminal helicase-binding domain. The C-terminal domain of DnaG (DnaG-CTD) interacts with the N-terminal domain of DnaB and stimulates the DnaB helicase activity (12). The stability of the interaction between DnaG primase and DnaB helicase varies among eubacteria. The helicase-primase interaction has been reported to be weak in *Escherichia coli* (13), moderate in *Helicobacter pylori* (14, 15) and strong in *Bacillus stearothermophilus* (9). Despite its relatively low resolution, the only available structure of the complex of DnaG-CTD with DnaB helicase in *Bacillus stearothermophilus* did provide insight into the helicase-primase interaction pattern (16). Since the helicase-binding domain of primases is poorly conserved, we expect the structures from different organisms to provide insight into the mode of interaction between DnaG and DnaB.

DNA replication in *Mycobacterium tuberculosis (Mt)* is not well understood. This intracellular pathogen is the main cause of human tuberculosis (17), and according to the recent World Health Organisation report (WHO report 2017), tuberculosis is the leading cause of death from a single infectious agent. This pathogen can, under favorable conditions, hide in a dormant state in the host for long periods of time without producing conspicuous disease symptoms (18). Specifically, *M. tuberculosis* progresses to a state of non-replicating persistence to lie dormant in the host for such long periods of time (19). Mutations generated

by error-prone DNA replication also enhance *M. tuberculosis* virulence and lead to antibiotic resistance (20). Part of our lack of understanding of DNA replication in *M. tuberculosis* arises from the homologs of some essential proteins for efficient DNA replication in *E. coli* having not been identified in mycobacteria. In addition, the regulators of DnaA activity (Hda in *E. coli*, YabA in *B. subtilis*), replicative helicase loader protein (DnaC in *E. coli*, DnaI in *B. subtilis*) and replication terminator proteins (Tus in *E. coli* and RTP in *B. subtilis*), have not been characterized in *M. tuberculosis* (21-24). Recently, Rv0004 was found to have a role in *M. tuberculosis* DNA replication; it interacts with DNA and the DnaB (25). The mechanism by which the replisome of *M. tuberculosis* is formed from the DnaG-DnaB complex and various accessory proteins is still not clear. *M. tuberculosis* DnaB (*Mt*DnaB) has recently been functionally characterized (26). The *Mt*dnaB gene encodes 874 amino acid residues, 416 of which (from Leu400 to Asn815) form an intein (27). An intein is an internal segment of amino acids that are excised from a protein precursor to generate a functional protein that splits off from the main protein during posttranslational maturation (27). The lack of structural data for the DnaB-DnaG complex and the incomplete structural information for the individual proteins are obstructing our efforts to understand the molecular details that reinforce this essential interaction in *M. tuberculosis*. However, the crystal structure of the NTD of *Mt*DnaB has been reported (28). To further study the mode of interaction between *Mt*DnaB and *Mt*DnaG, we determined the structure of the helicase-binding domain of *Mt*DnaG.

The *MtDnaG-CTD* structure was observed to differ from the other available DnaG-CTD structures. We used crystal structures of *Mt*DnaB-NTD reported earlier (28) and the currently determined structure of *Mt*DnaG-CTD to derive a model of helicase-primase interactions in *M. tuberculosis*, and provided further evidence for this model by performing in vitro experiments We identified single amino acid residues at the helicase-binding domain of *Mt*DnaG that, when mutated to alanine, resulted in proteins with altered interactions with DnaB helicase. Further analysis revealed the possibility of cross-species interaction of *Mt*DnaB with other DnaG in eubacteria. Overall, these results substantially helped us to understand the mechanism of helicase-primase interactions specifically in *M. tuberculosis* and expand our knowledge on the diversity of bacterial DNA replication strategies, some of which may be organism specific.

Materials and methods:

Cloning of expression constructs for heterologous expression:

Mycobacterium tuberculosis (H37rv) genomic DNA and specific primers (Table S1) were used for the polymerase chain reaction (PCR) to amplify the coding regions of *Mt*dnaG-CTD (477 base pairs) and *Mt*dnaB-NTD (591 base pairs). The amplified products of *Mt*dnaG-CTD and *Mt*dnaB-NTD were cloned into the Nhe1/Xho1 site of the expression vector pET21c (Novagen, Madison, WI, USA) and the Nde1/Xho1 site of the expression vector pET28b (Novagen, Madison, WI, USA), respectively, with each containing the C-terminal His6 tag. Similarly, dnaG-CTD from Vibrio cholerae (O395 strain) was cloned into the Nco1/Xho1 site of pET28b (Rehman thesis). DNA sequencing confirmed the sequences of all constructs. The cloning of *Hp*dnaG-CTD and *Ec*dnaG-CTD was performed previously in our laboratory (15, 29).

Site-directed mutagenesis

*Mt*DnaG-CTD mutants (*Mt*DnaG-CTD I605A and *Mt*DnaG-CTD F615A) were prepared by following a PCR-based method using the pET21c expression vector encoding for the wild-type (*Mt*DnaG-CTD) protein as a template. The amplification of the expression vector was performed using the KOD hot-start DNA polymerase (EMD Millipore), following the manufacturer guidelines, and specific pairs of primers (Table S1) were used for the introduction of the desired mutation. PCR products were treated with DpnI enzyme (New England Biolabs) and transformed into DH5α *E. coli* cells. The mutations were confirmed by performing DNA sequencing.

Overexpression and purification of recombinant proteins:

pET21c-*Mt*DnaG-CTD and point mutant (*Mt*DnaG-CTD I605A and *Mt*DnaG-CTD F615A) constructs were transformed into *E. coli* BL21 (DE3). For expression, 1% of the overnight-grown culture of a single colony was added to two liters of Luria-Bertani (LB) broth containing 100 μg/ml ampicillin, 1 mM glucose, and 1 mM $MgCl_2$. Cells were grown to an optical density of 0.6 (in 600 nm wavelength) at 37°C. Overexpression of protein was induced overnight at 16°C with 0.5 mM isopropyl-beta-D thiogalactopyranoside (IPTG). The cells were harvested at 8,000 rpm for 5 min and mixed with lysis buffer A (30 mM Tris-HCl [pH 7.5], 150 mM NaCl, 5 mM imidazole, 5 mM $MgCl_2$, 5 mM beta-mercaptoethanol [βME], 100 μM phenylmethylsulfonyl fluoride [PMSF]). After the addition of lysozyme (0.1 mg/ml), the cell suspension was incubated at 4°C for 30 min. The cell suspension was sonicated (Branson Ultrasonic Systems) in an ice-water mixture at 25% of the amplitude with

a pulse of 20 seconds each interspersed with an interval of 30 seconds, and cell lysate was treated with 0.1% Triton X-100, followed by incubation for 30 minutes on a rotating rocker at 4°C. The cell lysate was cleared by being subjected to centrifugation at 13,000 rpm for 45 minutes at 4°C and then purified using Ni-NTA Sepharose resin (GE Healthcare, Sweden) pre-equilibrated in buffer A. Proteins were eluted with buffer B (30 mM Tris-Cl [pH 7.5], 150 mM NaCl, 250 mM imidazole, 5 mM βME and 100 μM PMSF). The protein was then concentrated using an AmiconUltra-4 centrifugal filter ultracel-10K (Millipore, Billerica, MA, USA) and purified further by using a gel filtration chromatography Hi-Load Superdex-75 16/60 column (GE Healthcare) pre-equilibrated with buffer C (30 mM Tris-Cl [pH 7.5], 150 mM NaCl, 5% glycerol and 5 mM βME). Peak fractions were checked using SDS–12% PAGE and pooled together.

*Hp*DnaG-CTD and *Ec*DnaG-CTD were purified as described earlier (15, 29). *Mt*DnaB-NTD and *Vc*DnaG-CTD constructs were transformed into *E. coli* BL21 (DE3) cells. Cells containing plasmids were grown in LB media in the presence of either 50 μg/ml kanamycin (*Mt*DnaB-NTD) or 100 μg/ml ampicillin (*Vc*DnaG-CTD) up to an OD_{600} of 0.6-0.8 at 37°C. Overexpression of protein was induced overnight at 20°C with 0.2 mM IPTG. Bacterial cell pellets were re-suspended in lysis buffer A and cells were disrupted by subjecting them to sonication. The lysate was cleared by subjecting them to centrifugation at 13,000 g for 45 minutes at 4°C; and the proteins here were purified by being using Ni-NTA Sepharose beads pre-equilibrated in buffer A, and eluted with buffer B. Final purification was performed using gel filtration on a Hi-Load Superdex-75 16/60 column pre-equilibrated with buffer G (20 mM HEPES pH 7.4 and 150 mM NaCl). The homogeneity of protein was checked with (12%) SDS-PAGE.

Overexpression and purification of the selenomethionine-labeled *Mt*DnaG-CTD:

Selenomethionine (Se-Met)-labeled *Mt*DnaG-CTD was prepared to provide phases necessary for structure determination. Se-Met-labeled protein was purified under reducing conditions to prevent the formation of a mixture of reduced and oxidized Se-Met-labeled proteins. Protein labeling was carried out using media supplied by Molecular Dimensions. The concentration of selenomethionine was maintained at about 25 mg/liter. Initially, the primary culture of MtDnaG-CTD-transformed *E.coli* BL21 (DE3) was grown in LB medium overnight. The next morning, cells were harvested by centrifuging the overnight grown culture at 4,000 rpm for 5 minutes. The pellet obtained was resuspended in a purely selenomethionine medium and

the process was repeated once again to remove any traces of LB medium. After inoculation of the secondary culture in complete selenomethionine medium, the culture was allowed to grow for 4 hours until the OD_{600} reached 0.5 at 37°C. After this procedure, the cells were induced with 1 mM IPTG, and then shaken for another 8 hours at 30°C, followed by being harvested at 8,000 rpm for 3 minutes and stored at −80°C for downstream processing. The remaining steps of the purification were similar to those followed for the native *Mt*DnaG-CTD, as described above.

Crystallization of *Mt*DnaG-CTD:

Native *Mt*DnaG-CTD (6 mg/mL) and selenomethionine-labeled *Mt*DnaG-CTD (5 mg/mL) were subjected to extensive crystallization trials using the hanging drop method. Initially, the automated machine MOSQUITO™ (AIRF, JNU) was used to set hanging crystallization drops containing 500 nL protein solution plus 500 nL reservoir solution in 96-well Greiner-CELLSTAR plates for screening with Hampton and Molecular Dimensions crystallization screens. Small protein needles were grown using the reservoir solution of Hampton Peg Ion screen 25 (20% w/v PEG 3350, 0.2 M magnesium acetate tetrahydrate). The conditions were replicated and improved in 24-well Linbro plates by mixing 2 µl protein solution with 2 µl reservoir solution and equilibrating the resulting mixture against a 500 ml reservoir solution at 4°C. The best selenomethionine-labeled protein crystals were grown with a reservoir solution of 18% w/v PEG 3350, 0.15 M magnesium acetate and MOPS buffer pH 7.6. Crystals appeared after one week and grew to full size within two weeks.

Data collection, processing and structure determination:

The crystals were flash frozen in liquid nitrogen with a cryoprotectant solution containing 20% PEG 400 mixed with mother liquor (18% w/v PEG 3350, 0.15 M magnesium acetate and MOPS buffer with pH 7.6). The X-ray data for Se-Met-labeled crystals were collected at the BM14 synchrotron beamline (ESRF, Grenoble, France) at a selenium peak wavelength of 0.9786 Å. The data sets were indexed, integrated and scaled using HKL-2000 data processing software (30). The input diffraction data were prepared using programs of the CCP4 suite (31, 32). A partial structure was solved using the single wavelength anomalous scattering protocol of Auto-Rickshaw of the EMBL-Hamburg automated crystal structure determination platform ((33). Anomalous data were used to calculate FA values using the program SHELXC (34). Based on an initial analysis of the data, the maximum resolution for substructure determination and initial phase calculation was set to 2.0 Å. All of the three

heavy atoms expected (three selenomethionine atoms for each molecule of the asymmetric unit) were found using the program SHELXD (35). The correct hand for the substructure was determined using the programs ABS (36) and SHELXE (37). Initial phases were calculated after density modification using SHELXE. The initial phases were improved by carrying out density modification and phase extension to a resolution of 1.57 Å using the program RESOLVE (38). The initial model was built using the auto-build program ARP/wARP (39). The regions missing from this model were built manually using the COOT graphics package (40), and refinement of the resulting model was carried out with REFMAC5 (41). The structure was further improved after the iterative model building and refinement using COOT (42) and REFMAC5 (43), respectively. The final crystal structure was well refined with excellent electron density and validated by using PROCHECK (44) of the CCP4 suite. The figures presented here were generated using PyMOL (45). The data collection and refinement statistics are given in Table 1. Coordinate and structure factor files for the crystals of *Mt*DnaG-CTD have been deposited with the Protein Data Bank under accession code PDB ID 5Z51.

Construction of the helicase-primase complex model in *M. tuberculosis*:

The structure of *Mt*DnaB-NTD reported earlier by the Tsodikov group (28) and our current structure of *Mt*DnaG-CTD, as well as the structure of the *Bst*DnaB-*Bst*DnaG-CTD (helicase-primase) complex (16) as a template, were used to model the *Mt*DnaB-NTD/*Mt*DnaG-CTD (helicase-primase) complex of *M. tuberculosis*. The structure of *Mt*DnaB-NTD was superimposed on that of *Bst*DnaB-NTD, and the globular region (GR) of *Mt*DnaG-CTD was superimposed on the GR of *Bst*DnaG-CTD of the complex. To superimpose the helical hairpin region (HHR) of *Mt*DnaG-CTD with the HHR of the *Bst*DnaB-*Bst*DnaG-CTD complex, the HHR of *Mt*DnaG-CTD was rotated by about 32°, keeping the linker between the HHR and the GR as the center of rotation, to obtain the lowest root mean square deviation (RMSD) between the *Bst*DnaG-CTD and *Mt*DnaG-CTD structures. This modification maximized the interactions between primase and helicase. The structure of the modeled complex was relaxed, by carrying out molecular energy minimization using the GROMACS molecular dynamics package (46), in order to eliminate bad contacts. Similarly, we constructed models of three complexes of MtDnaB-NTD, one with *Ec*DnaG-CTD, and the other two with *Vc*DnaG-CTD and *Hp*DnaG-CTD, respectively, to study the possible interactions and stability levels of the heterocomplex.

Molecular dynamics simulation for binding affinity estimation:

Molecular dynamics simulations were performed on the models of the *Mt*DnaB-NTD complex with DnaG-CTD, DnaG-CTD I605A, and DnaG-CTD F615A, using the GROMACS 5.1.4 package, with the GROMOS 96 force field (47). To create the *in silico* mutants, isoleucine 605 and phenylalanine 615 were each changed to alanine using PyMOL (45) to make the two different DnaG-CTD mutants. These proteins were each solvated in a dodecahedron solvation box by using a simple point charge (SPC) (48) of water molecules. Na^+ and Cl^- ions were included to neutralize the system. The energy was minimized using 500 steps of steepest descent, followed by 1,000 steps of conjugate gradient. For maintaining the system in a stable environment (300 K, 1 bar), Berendsen temperature coupling (49) and Parrinello-Rahman pressure coupling (50) were employed and were set to 0.1 and 2.0 ps for temperature and pressure, respectively. A particle mesh Ewald (PME) algorithm (51) was run to measure the electrostatic and Van der Waals interactions; here, the cut-off distance for short-range VdW (rvdw) interactions was set to 1.4 nm, and Coulomb cut-off (r coulomb) and a neighbor list (rlist) was set to 0.9 nm. The LINCS algorithm (52) was used to measure all of the constrained bond lengths and the time step was set to 0.002 ps. The complexes in a medium were equilibrated for 100 ps in NPT and NVT ensembles. Finally, a 25 ns molecular dynamics simulation was carried out for each tested structure. All trajectories were stored every two ps for further analysis. Similar molecular dynamics simulation steps, as described above, were used on the three models of the complexes of *Mt*DnaB-NTD, i.e., with *Ec*DnaG-CTD, *Hp*DnaG-CTD, and *Vc*DnaG-CTD. The binding free energies were calculated by using the molecular mechanics Poisson-Boltzmann surface area (MM/GBSA) method as implemented in GROMACS tool (46). In our study, the free binding energy (ΔG) was calculated by using van der Waals (ΔE_{vdW}), electrostatic (ΔE_{elec}), polar solvation (ΔG_{polar}) and nonpolar ($\Delta G_{non\text{-}polar}$) energy contributions.

Biosensor binding studies:

To determine the difference between the binding affinity of *Mt*DnaB-NTD for *Mt*DnaG-CTD and those of *Mt*DnaB-NTD for the *Mt*DnaG-CTD mutants and other bacterial DnaG-CTDs, biosensor-based binding experiments were performed using an Autolab surface plasmon resonance (SPR) instrument at the Advanced Instrumentation Research Facility, Jawaharlal Nehru University, New Delhi, India. The surface of SPR chip (a self-assembled monolayer of 11-mercaptoundecanoic acid [MUA] on a gold surface; Autolab) was first activated with carbodiimide (EDC; 0.2 M) and *N*-hydroxysuccinimide (NHS; 0.05 M)/*N*-ethyl-*N*-(diethyl

aminopropyl). *Mt*DnaG-CTD, *Mt*DnaG-CTD mutants, and other bacterial DnaG-CTDs were immobilized (separately on different chips) onto the activated sensor surface at a concentration of 20 μg/ml in filtered (0.22-μm pore size) and degassed 10 mM sodium acetate buffer (pH 4.5). The chip had two channels: channel one was used for immobilization of ligand; and channel two was used as a blank (the signals of the analyte with a ligand-free surface). After ligand immobilization, the surface was blocked with 100 mM ethanolamine at pH 8.5, followed by regeneration using 1 M NaCl. The running buffer constituents were the same as those recommended for HBS BIAcore running buffer (10 mM HEPES [pH 7.4], 150 mM NaCl, 3 mM EDTA, 0.05% P-20 surfactant). The association kinetics for DnaG-CTDs was monitored for 400 seconds, followed by monitoring of the dissociation kinetics for the next 300 seconds. *Mt*DnaB-NTD samples of various concentrations were prepared in running buffer and injected at the rate of 20 μl/min across the sensor surface. The concentrations of *Mt*DnaB-NTD used against *Mt*DnaG-CTD were 1500, 1000, 750, 500, 250 and 125 nM, whereas, for the *Mt*DnaG-CTD mutants and other bacterial DnaG-CTDs, the concentrations used were 2000, 1500, 1,000, 750, 500 and 250 nM. Bovine serum albumin protein samples at various concentrations (2000, 1500, 1,000, 750, 500 and 250 nM) were used as negative controls. Signal changes on the activated/blocked control panel were subtracted from the DnaG-DnaB binding interactions using an inline reference signal, and the subtracted sensorgrams were analyzed. The surface was regenerated with a buffer consisting of two manually delivered pulses of 1 M NaCl. All of the data were recorded at 25°C. The data analysis was performed using Autolab SPR kinetic evaluation software.

Results:

The overall structure of MtDnaG-CTD:

To obtain a more detailed view of the interaction between the helicase and primase of *Mycobacterium tuberculosis*, we set out to determine a high-resolution crystal structure of the helicase-binding domain of *Mt*DnaG (*Mt*DnaG-CTD) and indeed did so, to a resolution of 1.58 Å (Fig. 2b). Molecular replacement failed to give any solution, indicative of the low homology and structural similarity with all of the previously determined DnaG-CTDs. Se-Met-labeled protein crystals were used to obtain anomalous data and experimental phases. The Se-Met-labeled protein crystal diffracted X-rays to 1.58 Å at BM14 of ESRF and belonged to the space group $P2_12_12$ with cell dimensions of a = 48.2 Å, b = 137.5 Å, c = 36.5 Å (α= β= γ = 90°). The final model yielded an R_{factor} of 18.1% and R_{free} of 20.8%, with good

electron density (Fig. 2a). One complete *Mt*DnaG-CTD (residues 481 to 637) and one helical hairpin cleaved from the *Mt*DnaG-CTD (residues 569 to 637) were identified per asymmetric unit (Fig. 2b) and these two protein chains were stabilized with a few van der Waals interactions (Fig. 2d). The overall CTD structure of *Mt*DnaG displayed two subdomains, the N-terminal globular region (GR) subdomain (residues 482 to 562) and C-terminal helical hairpin region (HHR) subdomain (residues 568 to 637), connected by a small loop (6 residues) (Fig. 2c). Inspection of the *Mt*DnaG-CTD structure showed it to lack any beta sheet and to consist instead primarily of alpha helices, specifically ten of them with the first eight forming the GR and the last two forming the HHR (Fig. 2c). The gel filtration chromatography profile of *Mt*DnaG-CTD in solution on a calibrated Superdex G75 16/60 column revealed a sharp elution peak at 70.3 ml, and an observed molecular weight of ~38 kDa, indicating *Mt*DnaG-CTD (~19 kDa) to exist in a dimeric state in solution (Fig. 1). Electron density corresponding to the GR or N-terminal residues (residues 481 to 568) of *Mt*DnaG-CTD was not visible in one monomer due to the cleavage of *Mt*DnaG-CTD. *Mt*DnaG-CTD cleaved into two fragments (GR and HHR), and the fragments were also identified using MALDI-TOF (Fig. S2). A total of 232 amino acid residues, one polyethylene glycol (PEG) molecule, one acetate molecule and 320 water molecules were determined to be present per asymmetric unit. A Ramachandran analysis showed 99.1% of the residues in the favored region, 0.5% in the generously allowed region, and 0.4% in the disallowed region.

The *Mt*DnaG-CTD structure differs from other DnaG-CTD structures in the HHR:

The 1.58-Å resolution of the current *M. tuberculosis* DnaG-CTD structure was higher than those of the other four available DnaG-CTD crystal structures, i.e., those determined for *E. coli* (resolution to 2.8 Å, PDB ID 1T3W), *B. stearothermophilus* (2.9 Å, 2R6A), *H. pylori* (1.7 Å, 4EHS) and *V. cholerae* (2.4 Å, 4IM9) (15, 16, 53). Structural alignments of *Mt*DnaG-CTD with these structures were done, using LSQMAN (54), to evaluate the differences between them. We specifically aligned the GRs of these structures to show the diversity in the relative orientation of the HHR. (Fig. 3). The RMSDs between *Mt*DnaG-CTD and *Ec*DnaG-CTD, *Bst*DnaGCTD, *Hp*DnaG-CTD, and *Vc*DnaG-CTD were found to be 1.89 Å (45 Cα atoms), 1.85 Å (71 Cα atoms), 1.59 Å (62 Cα atoms), and 1.77 Å (43 Cα atoms), respectively. However, DnaG-CTDs have been found to share less than 10% sequence identity (Fig. S1). The total structural alignments of *Mt*DnaG-CTD with *Ec*DnaG-CTD, *Bst*DnaGCTD, *Hp*DnaG-CTD, and *Vc*DnaG-CTD using RAPIDO (55) yielded RMSDs of 12.58 Å (101 Cα atoms), 6.81 Å (131 Cα atoms), 5.24 Å (108 Cα atoms) and 4.28 Å (122 Cα

atoms), respectively (Table S2). Inspection of the superpositions of these structures suggested the GR of *Mt*DnaG-CTD to be structurally more similar to that of *Hp*DnaG-CTD than to those of the other DnaG-CTDs. Overall, the arrangement of helices in *Mt*DnaG-CTD was observed to be similar to that in *Vc*DnaG-CTD (Table S2). The number of alpha helices present in each CTD structure was also found to differ: ten helices in the *Mt*DnaG-CTD structure, six each for *Hp*DnaG-CTD and *Vc*DnaG-CTD, and seven each for *Ec*DnaG-CTD and *Bst*DnaG-CTD (Fig. S4). These differences in the number of alpha helices resulted for the most part in offsets between the boundaries of the helices in the structure-based alignment (Fig. 3). Comparison of the *Mt*DnaG-CTD structure with other DnaG-CTD structures showed a larger deviation between the entire structures than between the GRs. The HHRs from the different sources adopted different orientations relative to the GRs and showed much greater structural differences than did the GRs. Moreover, the B factors of the HHRs and linkers were found to, in general, be greater than those of the GRs (Table S3), indicating relatively high flexibility levels for the HHR and linker regions. Also, note the lower overall B-factor of the *Mt*DnaG-CTD structure than of the other known CTD structures (Table S3). Interestingly, despite all of the DnaG-CTD structures consisting of a GR as well as HHR, multiple sequence alignment (MSA) of amino acid sequences (Fig. S1b) and structural alignment did not yield any significant homology.

_Mt_DnaG-CTD interacts at the dimer-dimer interface of _Mt_DnaB-NTD using mostly hydrophobic residues of the HHR:

A previous investigation showed *Mt*DnaB-NTD (residues 21–197, including the linker) to form a stable dimer in solution, and a truncated form of the protein (residues 21–134) to be monomeric, indicating the region spanning residues 135–197 to be required for dimer stability (28). Therefore, we also generated these monomeric and dimeric *Mt*DnaB-NTD constructs, using the same method as described earlier (28), for our studies. Specifically, biosensor binding studies of each of these two constructs with *Mt*DnaG-CTD was carried out using surface plasmon resonance by passing the respective *Mt*DnaB-NTD construct on immobilized *Mt*DnaG-CTD. The results of these two experiments showed *Mt*DnaG-CTD to not interact with the monomeric form of *Mt*DnaB-NTD (Fig. S3b) but to do so with the dimeric form (Fig. S3c).

The lack of a high-resolution structure of the DnaB-DnaG complex and the fragmented structural information for the individual proteins have impeded our attaining a detailed understanding of this crucial binary protein interaction. The only complex structure of the

active hexameric *Bst*DnaB (PDB: 2R6A) was determined as a trimer of dimer and three *Bst*DnaG-CTDs were found at the dimer-dimer interface of *Bst*DnaB (16). Similarly, the crystal structure of the *Mt*DnaB-NTD (PDB: 2R5U) has been reported to form a hexameric ring having two distinct interfaces; an extensive hydrobhobic interface, stabilizing a dimer (monomer-monomer) of *Mt*DnaB-NTD, and other less extensive interface formed between the dimers, connecting three of them into a hexamer (28). The dimer-dimer interface of *Mt*DnaB-NTD was considered as a DnaG binding site (28). Here, we produced a model of the *M. tuberculosis* DnaB-DnaG complex by one molecule of *Mt*DnaG-CTD interacted to two molecules of *Mt*DnaB-NTD, involved at the dimer- dimer interface similar to BstDnaB-DnaG-CTD complex structure (16). Superimpositions of the GR of the *Mt*DnaG-CTD crystal structure with GRs from the crystal structures of *Ec*DnaG-CTD, *Vc*DnaG-CTD, BstDnaG-CTD, and *Hp*DnaG-CTD respectively (Fig. 3), showed the orientation of the HHR to vary considerably between the different species and suggested the flexibility of the HHR. Inspection of the HHR in the unbound form of *Bst*DnaG-CTD showed it to be oriented 65° away from its orientation in the helicase-bound form (Fig. S5a), and this difference was attributed to the flexibility of the HHR. Therefore, the HHR of *Mt*DnaG-CTD was rotated by 32° to model the *Mt*DnaG-CTD/*Mt*DnaB-NTD complex using the *Bst*DnaB/DnaG-CTD (PDB: 2R6A) complex structure as a template (Fig. S5). A molecular dynamics simulation of the model was carried out for 25 nanoseconds and revealed the stability of the model after the first few nanoseconds of the simulation (Fig. 4c). When producing our model of the *Mt*DnaG-CTD/*Mt*DnaB-NTD complex, we maximized the number of interactions between the HHR of *Mt*DnaG-CTD and *Mt*DnaB-NTD (Fig. 4a). DnaG mutations in the part of the extreme C-terminal region, containing the HHR, were shown to be disruptive to the DnaB–DnaG complex (56), hence indicating this part of the C-terminal region to be crucial for the helicase-primase interaction. Our modeling highlighted the amino acid residues present in the helicase-primase complex interface that may participate in these crucial interactions and stabilize the complex (Fig. 4a). The low homology between the DnaG-CTD sequences (Fig. S1a) together with the presence of different yet consistently hydrophobic amino acid residues at the *M. tuberculosis* (Fig. 4a, 5b) and *B. stearothermophilus* (16) helicase-primase interface suggests the involvement of hydrophobic interactions in stabilizing the helicase-primase complex. According to the model, I605 and F615 from the primase of *M. tuberculosis* make extensive hydrophobic interactions with helicase, and these residues, in particular, may be playing a crucial role in stabilizing the helicase-primase complex (Fig. 4b).

Ile605 in the HHR is crucial for helicase-primase interaction in *Mycobacterium tuberculosis*:

To elucidate the main drivers of the helicase-primase association and to identify hotspots, we turned to quantify individual contributions of potential key residues, guided by a structural analysis of our proposed *M. tuberculosis* DnaG-CTD/DnaB-NTD complex model. To accomplish this goal, we mutated a couple of crucial contact residues modeled to be present at the interface of the *M. tuberculosis* DnaG-DnaB complex. Inspection of our model (Fig. 4a) indicated a large non-conserved hydrophobic residue, Ile605, to be the most prominent surface-accessible hydrophobic residue on the HHR of *Mt*DnaG-CTD (Fig. 4b). Inspection of our model indicated Ile605, of the loop connecting αH9 and αH10 of the HHR, to be stabilizing the complex. Ile605 present at the crucial position, flexible loop within the HHR, through that the whole HHR get hooked by this residue. Our analysis also indicated Phe615, another large hydrophobic residue of *Mt*DnaG-CTD, to be present at the *M. tuberculosis* DnaG-DnaB interface (Fig. 4b). This residue, being present at helix 10 of *Mt*DnaG-CTD, was observed to be less accessible than *Mt*DnaG-CTD Ile605 present at the connecting loop of the HHR. The stability levels of the helicase (*Mt*DnaB-NTD) in complex with native *Mt*DnaG-CTD, mutant *Mt*DnaG-CTD F615A, and mutant *Mt*DnaG-CTD I605A were analyzed by carrying out molecular dynamics simulations (Fig. 4c). Furthermore, the binding energy levels of all of the complex models were calculated using the MM/GBSA method of GROMACS (Table S4). The binding energy values of helicase (*Mt*DnaB-NTD) in complex with native *Mt*DnaG-CTD, mutant *Mt*DnaG-CTD F615A, and mutant *Mt*DnaG-CTD I605A were -100 kjoule/mol, -75.3 kjoule/mol and -60 kjoule/mol, respectively (Table S4). The binding affinities for the native and mutant *Mt*DnaG-CTDs in complex with *Mt*DnaB-NTD were also measured *in vitro* using surface plasmon resonance (SPR) spectroscopy (Fig. 5). Native and mutant *Mt*DnaG-CTDs were immobilized on three different chips under identical conditions, and different concentrations of *Mt*DnaB-NTD were passed as an analyte with HBS buffer over the immobilized proteins. The K_D values of MtDnaB-NTD in complex with MtDnaG-CTD, MtDnaG-CTD F615A, and MtDnaG-CTD I605A were measured to be 370 nM, 444 nM, and 2.8 µM, respectively (Fig. 5). The SPR results indicated a 10-fold decrease in the binding affinity of *Mt*DnaB-NTD with *Mt*DnaG-CTD I605A compared to that of *Mt*DnaB-NTD with native *Mt*DnaG-CTD. Although the measured difference in binding affinity to the native (*Mt*DnaG-CTD) and mutated (*Mt*DnaG-CTD I605A) proteins was relatively large considering that only a single residue was mutated, there was still observed significant affinity between the helicase and the mutated (*Mt*DnaG-CTD I605A) primase. In

our current study, we noticed the important role of the loop connecting the two helices of the HHR in stabilizing the helicase-primase complex. This major role of the loop within the HHR in the stability of the helicase-primase complex was indicated despite no common amino acid residues present at the same position of different DnaG-CTDs. Both *in silico* calculations and *in vitro* experiments (SPR) suggested a significant role of Ile605 of *Mt*DnaG in the stabilization of the helicase-primase (*Mt*DnaB/*Mt*DnaG) complex in *M. tuberculosis*.

Mycobacterium DnaB-NTD interacts to DnaG-CTD of other organisms with low binding affinity:

The helicase-binding domain of DnaG and primase-binding region of DnaB are the least conserved regions of their respective proteins. Despite the insignificant homology between the various DnaG-CTD sequences, the DnaG-CTD structures were found to be similar, especially in their all having globular and HHR regions. A functional complementation of *Ec*DnaB with *Hp*DnaB has been reported earlier (57), indicative of an interaction between *Hp*DnaB with *Ec*DnaG. To further test whether the cross-species helicase-primase interaction takes place even with such a low sequence similarity, we performed a study of the interaction of *Mt*DnaB-NTD with other DnaG-CTDs. We studied the specificity in the interaction of DnaB with non-cognate DnaG by generating an *in silico* model of the complex and further tested the biophysical interactions of *Mycobacterium* DnaB-NTD with other DnaG-CTDs, i.e., from *E. coli*, *H. pylori,* and *V. cholerae*. The steps used to *in silico* build and analyze the models of *Mt*DnaB-NTD in complex with DnaG-CTDs from *E. coli*, *H. pylori,* and *V. cholerae* (Fig. 6) were similar to the strategy used in producing the model of the *M. tuberculosis* helicase-primase complex. The free energy values of the binding of native *Mt*DnaB-NTD to *Ec*DnaG-CTD, *Hp*DnaG-CTD, and *Vc*DnaG-CTD were -13.7 kJ/mol, -71.5 kJ/mol and -58.4 kJ/mol respectively. The *in silico* studies suggested the *Mt*DnaB-NTD/*Hp*DnaG-CTD complex to be more stable than the *Mt*DnaB-NTD/*Ec*DnaG-CTD and *Mt*DnaB-NTD/*Vc*DnaG-CTD complexes. We also carried out biosensor binding studies using surface plasmon resonance (SPR) spectroscopy. Specifically, the binding affinities of native *Mt*DnaB-NTD for *Ec*DnaG-CTD, *Hp*DnaG-CTD, and *Vc*DnaG-CTD were measured *in vitro* using SPR spectroscopy. The K_D values for the complexes of *Mt*DnaB-NTD with *Ec*DnaG-CTD (Fig. 7a) and *Hp*DnaG-CTD (Fig. 7b) were determined from these studies to be 7.4 µM and 3.4 µM, respectively, while *Mt*DnaB-NTD formed an unstable association with *Vc*DnaG-CTD (Fig. 7c). However, the cross species helicase-primase binding did not show the order

as predicted by the *in silico* studies. But both *in silico* as well as *in vitro* cross- study on *Mt*DnaB-NTD to DnaG-CTD of other bacteria showed the significant decrease in binding affinities compared to the binding affinity of native *Mt*DnaG-CTD and *Mt*DnaB-NTD interaction. Moreover, the decrease in the binding affinity of DnaG to DnaB might affect the activity of these proteins during the replication process.

Discussion

In this study, the crystal structure of the DnaB helicase-binding domain of DnaG primase (*Mt*DnaG-CTD) of *Mycobacterium tuberculosis* was determined to gain insight into the helicase-primase interaction, crucial for the survival of eubacteria. The complex between the DnaG primase and the DnaB helicase unwinds duplex DNA at the eubacterial replication fork and synthesizes the RNA primers required for initiation of DNA replication (58, 59). An alignment of DnaG-CTD sequences showed high diversity. Therefore, the *Mt*DnaG-CTD crystal structure could not be solved using molecular replacement, and we instead used SAD phasing. Analysis of the determined structure highlighted the flexibility of the DnaG-CTD scaffold, which has been observed in other DnaG-CTDs, and supported the existence of hinge bending within the protein. The hinge point located in the linker between the globular region (GR) and hairpin region (HHR) apparently allows these two regions of the protein to become close together or far apart, resulting in an alternation between elongated and contracted shapes. Also, the flexibility of the helicase-binding domain has been shown to be due in large part to the identified hinge point, and to be the main factor maximizing the interactions with DnaB helicase. A comparison of *Mt*DnaG-CTD with other DnaG-CTD structures suggested some degree of conservation of the protein fold, in particular in the GR. Overall; the HHR appears to have the major role in the interaction with DnaB and shows a high degree of diversity in eubacteria.

We also constructed *in silico* a hypothetical model of the *Mt*DnaB-NTD/*Mt*DnaG-CTD complex, based on the *Bst*DnaB/DnaG-CTD complex crystal structure (16), and found the hypothetical model to be consistent with previous mutagenesis studies (56). In the model of the *Mt*DnaB-NTD/*Mt*DnaG-CTD complex, the HHR of *Mt*DnaG-CTD covers the interface otherwise making dimer-dimer interactions in the hexameric form *Mt*DnaB-NTD. Also, Ile605 and Phe615 are large non-conserved hydrophobic residues present at the HHR of *Mt*DnaG and, according to our model, interact with *Mt*DnaB (Fig. 4a). Our analysis also indicated Ile605, a hydrophobic residue present on the loop of the HHR of *Mt*DnaG-CTD, to interact with Val82 and Ala85 of *Mt*DnaB-NTD and hence to be crucial for the formation of

the complex. Our *in silico* study of the *Mt*DnaB/*Mt*DnaG complex also indicated a decrease in binding energy upon mutation of *Mt*DnaG-CTD Ile605 to alanine as well as upon the mutation of its Phe615 to alanine (Table S4). To provide further evidence for the results of these binding affinity calculations, we also carried out *in vitro* studies involving expressing both mutants of *Mt*DnaG-CTD and using SPR spectroscopy to measure their affinities for DnaB-NTD. In this biosensor binding study, the affinity of the Ile605Ala *MT*DnaG-CTD mutant for *Mt*DnaB-NTD was shown to be ten times weaker than the affinity of the native *Mt*DnaG-CTD for *Mt*DnaB-NTD. These findings helped us validate the model and identify Ile605 as one of the key *Mt*DnaG residues stabilizing the association of *Mt*DnaG with *Mt*DnaB.

Analysis of our current structure, biosensor results and the results of the previous studies of DnaB-DnaG complexes (15, 16) led us to conclude that the loop connecting the two helices of the HHR is apparently largely responsible for the stability of this complex. The sequence and structural analysis indicated the non-conserved residue Ile605, present at the helical hairpin loop of DnaG, to be specific to *M. tuberculosis* and to be a key residue for the interaction between DnaB and DnaG (Fig. 4). The earlier study in *H. pylori*, residue Phe537 of the helical hairpin loop of DnaG-CTD has been shown to play a key role in the stability of the DnaB-DnaG complex (15). The observation of amino acid residues present at the helical hairpin loop of DnaG-CTDs suggested the presence of non-conserved hydrophobic residues that play a significant role in the stability of the DnaB-DnaG complex. The flexibility of the loop apparently allows for maximizing the access of its hydrophobic residues for contacts with DnaB-NTD.

The binding affinity of *Mt*DnaB for *Mt*DnaG was measured to be in the nanomolar range while the binding affinities of *Mt*DnaB with other DnaGs from different species were measured in the micromolar range (Fig. 5, Fig. 7). Such a relatively low level of heterologous binding affinity might be expected to result in relatively poor helicase and primase activities in the case of complementarity with other organisms. A suboptimal activity level of any DNA replication initiation protein would, in turn, be expected to slow down replication fork progression (2). Though the helicase of *H. pylori* showed functional complementarity with *E. coli* (57), such complementarity of this helicase seems to be specific to just a few organisms. Therefore, *Mt*DnaB may show complementarity with *E. coli* and *H. pylori* but not in *V. cholerae*. The pattern of interactions between DnaG and DnaB seems to vary from one organism to the next, and therefore any change in the affinity between these proteins may affect not only the initiation of replication but also overall DNA replication.

Based on the crystal structure and our model, we can suggest a possible mechanism to explain the regulation of DnaB-DnaG activity at the replication fork. Bacterial primases primarily adopt a compact state but can transition to an extended conformation (60). In the unbound form or free form of DnaG, the HHR may fold back to interfere with the RNA polymerase domain or the Zn-binding domain. According to the proposed mechanism, the moment when DnaG-CTD comes into contact to helicase the GR side of DnaG-CTD might interact to the one molecule of *Mt*DnaB-NTD and this binding might enforce for the changing the orientation of HHR from the unbound state of DnaG. The HHR would bend in such a way as to properly bind to the second molecule of DnaB-NTD with comparatively stronger than first one, and the bending of the HHR away from the RNA polymerase domain or Zn-binding domain may lead to the activation of primase by coming into contact with DnaB. Because DnaG cannot effectively function as an independent molecule, i.e., unbound from DnaB, this mechanism would also provide a means to prevent unwanted priming at non-replicative sites in the cell. Future studies to further address this model in the context of the replisome will highlight how interactions between DnaG and the rest of the replisome synergize to regulate replication of the lagging strand.

In conclusion, the structural and biophysical data provided by our work have shed new light on the structural assembly and mode of binding of DnaG-CTD with DnaB-NTD in *M. tuberculosis* and unveiled mechanistic details and selectivity determinants for other eubacteria. The non-conserved interacting domains of DnaG-DnaB complex expand the study on the aspects of the species-specific evolution of the replication system.

Acknowledgments:

S.G. acknowledges DBT, DST-PURSE, and UGC RNW, University Grant Commission (SAP), D.P.S. acknowledges UGC for the fellowship, and R.V. acknowledges DST-SERB for the fellowship. We thank the Advanced Instrumentation Research Facility (AIRF) and JNU for providing the SPR facility. We also thank the staffs of BM14, ESRF for helping us with the collection of high-resolution X-ray data.

Author contributions:

SG and DPS conceived the idea and designed the experiments. DPS solved the *Mt*DnaG-CTD structure. SG, DPS and RV analyzed the structures and constructed the models. DPS

generated SDM and performed all of the biosensor studies. RV performed the MD simulations of all of the models and SAAR performed cloning. DPS wrote the manuscript taking input from SG, and SG reviewed the manuscript.

Abbreviations:

Mt, *Mycobacterium tuberculosis*; *Ec*, *Escherichia coli*; *Hp*, *Helicobacter pylori*; *Bst*, *Bacillus stearothermophilus*; *Vc*, *Vibrio cholerae*; DnaG-CTD, DnaG C-terminal domain; DnaB-NTD, DnaB N-terminal domain; RMSD, root mean square deviation;

References:

1. Kornberg A, Baker TA. DNA replication: Wh Freeman San Francisco; 1980.
2. Mott ML, Berger JM. DNA replication initiation: mechanisms and regulation in bacteria. Nature Reviews Microbiology. 2007;5(5):343.
3. Fang L, Davey MJ, O'Donnell M. Replisome assembly at oriC, the replication origin of E. coli, reveals an explanation for initiation sites outside an origin. Molecular cell. 1999;4(4):541-53.
4. LeBowitz J, McMacken R. The Escherichia coli dnaB replication protein is a DNA helicase. Journal of Biological Chemistry. 1986;261(10):4738-48.
5. Hacker KJ, Johnson KA. A hexameric helicase encircles one DNA strand and excludes the other during DNA unwinding. Biochemistry. 1997;36(46):14080-7.
6. O'Donnell ME, Li H. The ring-shaped hexameric helicases that function at DNA replication forks. Nature structural & molecular biology. 2018:1.
7. Corn JE, Berger JM. Regulation of bacterial priming and daughter strand synthesis through helicase-primase interactions. Nucleic acids research. 2006;34(15):4082-8.
8. Van Eijk E, Paschalis V, Green M, Friggen AH, Larson MA, Spriggs K, et al. Primase is required for helicase activity and helicase alters the specificity of primase in the enteropathogen Clostridium difficile. Open biology. 2016;6(12):160272.
9. Bird LE, Pan H, Soultanas P, Wigley DB. Mapping protein− protein interactions within a stable complex of DNA primase and DnaB helicase from Bacillus stearothermophilus. Biochemistry. 2000;39(1):171-82.
10. Tougu K, Marians KJ. The interaction between helicase and primase sets the replication fork clock. Journal of Biological Chemistry. 1996;271(35):21398-405.

11. Tougu K, Peng H, Marians KJ. Identification of a domain of Escherichia coli primase required for functional interaction with the DnaB helicase at the replication fork. Journal of Biological Chemistry. 1994;269(6):4675-82.
12. Davey MJ, Jeruzalmi D, Kuriyan J, O'Donnell M. Motors and switches: AAA+ machines within the replisome. Nature reviews Molecular cell biology. 2002;3(11):826.
13. Mitkova AV, Khopde SM, Biswas SB. Mechanism and stoichiometry of interaction of DnaG primase with DnaB helicase of Escherichia coli in RNA primer synthesis. The Journal of biological chemistry. 2003;278(52):52253-61.
14. Kashav T, Nitharwal R, Abdulrehman SA, Gabdoulkhakov A, Saenger W, Dhar SK, et al. Three-dimensional structure of N-terminal domain of DnaB helicase and helicase-primase interactions in Helicobacter pylori. PloS one. 2009;4(10):e7515.
15. Rehman SAA, Verma V, Mazumder M, Dhar SK, Gourinath S. Crystal structure and mode of helicase binding of the C-terminal domain of primase from Helicobacter pylori. Journal of bacteriology. 2013;195(12):2826-38.
16. Bailey S, Eliason WK, Steitz TA. Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase. Science. 2007;318(5849):459-63.
17. Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, et al. Tuberculosis. Nature Reviews Disease Primers. 2016;2:16076.
18. Grosset J. Mycobacterium tuberculosis in the extracellular compartment: an underestimated adversary. Antimicrobial agents and chemotherapy. 2003;47(3):833-6.
19. Wayne LG, Hayes LG. An in vitro model for sequential study of shiftdown of Mycobacterium tuberculosis through two stages of nonreplicating persistence. Infection and immunity. 1996;64(6):2062-9.
20. Boshoff HI, Reed MB, Barry CE, Mizrahi V. DnaE2 polymerase contributes to in vivo survival and the emergence of drug resistance in Mycobacterium tuberculosis. Cell. 2003;113(2):183-93.
21. Katayama T, Ozaki S, Keyamura K, Fujimitsu K. Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. Nature reviews Microbiology. 2010;8(3):163-70.
22. Bell SP, Kaguni JM. Helicase loading at chromosomal origins of replication. Cold Spring Harbor perspectives in biology. 2013;5(6):a010124.
23. Velten M, McGovern S, Marsin S, Ehrlich SD, Noirot P, Polard P. A two-protein strategy for the functional loading of a cellular replicative DNA helicase. Molecular cell. 2003;11(4):1009-20.
24. Brézellec P, Vallet-Gely I, Possoz C, Quevillon-Cheruel S, Ferat J-L. DciA is an ancestral replicative helicase operator essential for bacterial replication initiation. Nature communications. 2016;7:13271.

628 25. Mann KM, Huang DL, Hooppaw AJ, Logsdon MM, Richardson K, Lee HJ, et al. Rv0004 is a new
629 essential member of the mycobacterial DNA replication machinery. PLoS genetics.
630 2017;13(11):e1007115.
631 26. Zhang H, Zhang Z, Yang J, He ZG. Functional characterization of DnaB helicase and its
632 modulation by single-stranded DNA binding protein in Mycobacterium tuberculosis. The FEBS
633 journal. 2014;281(4):1256-66.
634 27. Yamamoto K, Low B, Rutherford SA, Rajagopalan M, Madiraju MV. The Mycobacterium
635 avium-intracellulare complex dnaB locus and protein intein splicing. Biochemical and biophysical
636 research communications. 2001;280(3):898-903.
637 28. Biswas T, Tsodikov OV. Hexameric ring structure of the N-terminal domain of
638 Mycobacterium tuberculosis DnaB helicase. The FEBS journal. 2008;275(12):3064-71.
639 29. Syson K, Thirlway J, Hounslow AM, Soultanas P, Waltho JP. Solution structure of the helicase-
640 interaction domain of the primase DnaG: a model for helicase activation. Structure. 2005;13(4):609-
641 16.
642 30. Otwinowski Z, Minor W. [20] Processing of X-ray diffraction data collected in oscillation
643 mode. Methods in enzymology. 276: Elsevier; 1997. p. 307-26.
644 31. Collaborative CP. The CCP4 suite: programs for protein crystallography. Acta
645 crystallographica Section D, Biological crystallography. 1994;50(Pt 5):760.
646 32. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4
647 suite and current developments. Acta Crystallographica Section D: Biological Crystallography.
648 2011;67(4):235-42.
649 33. Panjikar S, Parthasarathy V, Lamzin VS, Weiss MS, Tucker PA. Auto-Rickshaw: an automated
650 crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction
651 experiment. Acta Crystallographica Section D: Biological Crystallography. 2005;61(4):449-57.
652 34. Sheldrick GM. Experimental phasing with SHELXC/D/E: combining chain tracing with density
653 modification. Acta Crystallographica Section D: Biological Crystallography. 2010;66(4):479-85.
654 35. Schneider TR, Sheldrick GM. Substructure solution with SHELXD. Acta Crystallographica
655 Section D: Biological Crystallography. 2002;58(10):1772-9.
656 36. Hao Q. ABS: a program to determine absolute configuration and evaluate anomalous
657 scatterer substructure. Journal of applied crystallography. 2004;37(3):498-9.
658 37. Sheldrick GM. Macromolecular phasing with SHELXE. Zeitschrift für Kristallographie-
659 Crystalline Materials. 2002;217(12):644-50.
660 38. Terwilliger TC. Maximum-likelihood density modification. Acta Crystallographica Section D:
661 Biological Crystallography. 2000;56(8):965-72.

662 39. Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative
663 structure refinement. Nature Structural and Molecular Biology. 1999;6(5):458.
664 40. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta
665 Crystallographica Section D: Biological Crystallography. 2004;60(12):2126-32.
666 41. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the
667 maximum-likelihood method. Acta Crystallographica Section D: Biological Crystallography.
668 1997;53(3):240-55.
669 42. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta
670 Crystallographica Section D: Biological Crystallography. 2010;66(4):486-501.
671 43. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. REFMAC5
672 for the refinement of macromolecular crystal structures. Acta Crystallographica Section D: Biological
673 Crystallography. 2011;67(4):355-67.
674 44. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the
675 stereochemical quality of protein structures. Journal of applied crystallography. 1993;26(2):283-91.
676 45. DeLano WL. The PyMOL molecular graphics system. http://pymol org. 2002.
677 46. Kumari R, Kumar R, Consortium OSDD, Lynn A. g_mmpbsa ▪ A GROMACS tool for high-
678 throughput MM-PBSA calculations. Journal of chemical information and modeling. 2014;54(7):1951-
679 62.
680 47. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High
681 performance molecular simulations through multi-level parallelism from laptops to supercomputers.
682 SoftwareX. 2015;1:19-25.
683 48. Wu Y, Tepper HL, Voth GA. Flexible simple point-charge water model with improved liquid-
684 state properties. The Journal of chemical physics. 2006;124(2):024503.
685 49. Berendsen HJ, Postma Jv, van Gunsteren WF, DiNola A, Haak J. Molecular dynamics with
686 coupling to an external bath. The Journal of chemical physics. 1984;81(8):3684-90.
687 50. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular
688 dynamics method. Journal of Applied physics. 1981;52(12):7182-90.
689 51. Darden T, York D, Pedersen L. Particle mesh Ewald: An N· log (N) method for Ewald sums in
690 large systems. The Journal of chemical physics. 1993;98(12):10089-92.
691 52. Hess B, Bekker H, Berendsen HJ, Fraaije JG. LINCS: a linear constraint solver for molecular
692 simulations. Journal of computational chemistry. 1997;18(12):1463-72.
693 53. Oakley AJ, Loscha KV, Schaeffer PM, Liepinsh E, Pintacuda G, Wilce MC, et al. Crystal and
694 Solution Structures of the Helicase-binding Domain of Escherichia coli Primase o. Journal of
695 Biological Chemistry. 2005;280(12):11495-504.

54. Armougom F, Moretti S, Keduas V, Notredame C. The iRMSD: a local measure of sequence alignment accuracy using structural information. Bioinformatics. 2006;22(14):e35-e9.
55. Hasegawa H, Holm L. Advances and pitfalls of protein structural alignment. Current opinion in structural biology. 2009;19(3):341-8.
56. Tougu K, Marians KJ. The extreme C terminus of primase is required for interaction with DnaB at the replication fork. Journal of Biological Chemistry. 1996;271(35):21391-7.
57. Soni RK, Mehra P, Mukhopadhyay G, Dhar SK. Helicobacter pylori DnaB helicase can bypass Escherichia coli DnaC function in vivo. Biochemical Journal. 2005;389(2):541-8.
58. Ha T, Rasnik I, Cheng W, Babcock HP, Gauss GH, Lohman TM, et al. Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase. Nature. 2002;419(6907):638.
59. Lu Y-B, Ratnakar PV, Mohanty BK, Bastia D. Direct physical interaction between DnaG primase and DnaB helicase of Escherichia coli is necessary for optimal synthesis of primer RNA. Proceedings of the National Academy of Sciences. 1996;93(23):12902-7.
60. Corn JE, Pease PJ, Hura GL, Berger JM. Crosstalk between primase subunits can act to regulate primer synthesis in trans. Molecular cell. 2005;20(3):391-401.

Table 1: Crystallographic data and refinement statistics:

Crystallographic data	
X-ray source	BM14, ESRF, France
Wavelength (Å)	0.97
Space group	$P2_12_12$
Unit cell parameters	
a, b, c (Å)	48.28, 137.55, 36.51
α, β, γ (º)	90, 90, 90
Resolution range (Å)	68.78 -1.58
Completeness (%)	92.8 (99.8)
Total no. of observations	150161
No. of unique observations	40035
Redundancy	3.8 (3.9)
Average I/σ (I)	19.66 (1.94)
CC (1/2)	0.92 (0.72)
$R_{r.i.m.}$	0.088
Crystal mosaicity (º)	0.76
Refinement	
R-factor (%)	18.14
Free R-factor (%)	20.88
Mean B factor ($Å^2$)	16.1
Number of atoms	
Protein/PEG/acetate/water	1812/17/7/320
RMS deviation	
Bond length (Å)	0.017
Bond angles (°)	1.64
Ramachandran plot (%)	
In preferred region	99.1
In allowed region	0.5
Outliers	0.4
PDB ID	5Z51

Figure legends:

Fig. 1: *Mt*DnaG-CTD is a dimer in solution. (a) Size-exclusion chromatograph of the C-terminal domain of primase (*Mt*DnaG-CTD) passed through a HiLoad 16/60 Superdex 75 column. Ten-milliliter fractions were collected. The elution volume (70.3 ml) and the elution pattern of the protein are displayed. (b) The molecular mass of the eluted DnaG-CTD was deduced from the standard plot to be about ~ 35kDa, corresponding to the dimeric state of the protein. (c) SDS-PAGE showing fractions purified by gel filtration. The proteins were separated on SDS–12% PAGE and stained with Coomassie brilliant blue. Lane M shows the molecular markers; lane P shows the gel filtration fraction.

Fig. 2: Crystal structure of *Mt*DnaG-CTD. (a) A portion of the molecule shown with a 2Fo-Fc electron density map contoured at 1σ within 1.5 Å of selected atoms. (b) Cartoon diagram showing the crystal structure of the *Mt*DnaG primase CTD. The overall structure was observed to have one globular region connected to the helical hairpin region by a flexible loop. Chain A is colored purple, and chain B is cyan. One complete *Mt*DnaG-CTD (chain A) and one helical hairpin cleaved from this CTD (chain B) were identified per asymmetric unit. (c) Cylindrical representation of the MtDnaG-CTD crystal structure showing the arrangement of the ten helices connected by loops. Helices αH1 - αH8 formed the globular region and helices αH9 – αH10 formed the helical hairpin region. (d) Amino acid residues showed the hydrophobic interaction between the chain A and chain B in the crystal structure of *Mt*DnaG-CTD.

Fig. 3: Superpositions of the globular region of *Mt*DnaG-CTD with the globular regions of other DnaG-CTDs provided insight into the differences in the shapes and orientations of HHRs. Superposition of the globular region of the *Mt*DnaG-CTD (purple) crystal structure with the globular region of each of (a) *Ec*DnaG-CTD (green), (b) *Vc*DnaG-CTD (cyan), (c) *Bst*DnaG-CTD (violet) and (d)*Hp*DnaG-CTD (wheat). The helical hairpin regions (HHRs) of the different structures showed different orientations.

Fig. 4: Model of the *M. tuberculosis* helicase-primase complex. (a) Model of the *M. tuberculosis* helicase-primase complex. Experimentally determined structures of *Mt*DnaB-NTD (PDB: 2R5U) and *Mt*DnaG-CTD (PDB: 5Z51), as well as that of the *Bst*DnaB-*Bst*DnaG-CTD complex (PDB: 2R6A), were used to generate the model of the complex. In the complex model, two molecules of DnaB-NTD interacts with one molecule of DnaG-CTD.

The model was further relaxed to eliminate bad contacts using molecular energy minimization with the help of the GROMACS molecular dynamics package. (b) The crucial residues of the helical hairpin showing hydrophobic interactions stabilizing the helicase-primase complex. (c) RMSD analysis of the trajectories of native and mutants primase models. Time-dependent Cα RMSDs for all residues of DnaB-NTD complexed with the native DnaG-CTD (black), I605A mutant DnaG-CTD (red), and F615A mutant DnaG-CTD (green) showed structural convergence from the initial time point. No effect of the mutations on the DnaG-CTD structure was observed, but a significant effect on the binding affinity of the helicase for the primase was found.

Fig. 5: SPR sensorgram for measuring the binding affinity of the helicase for the primase. Sensorgrams showing the binding pattern of *Mt*DnaB-NTD at various concentrations (legends) with immobilized (a) native *Mt*DnaG-CTD, (b) F615A mutant *Mt*DnaG-CTD, and (c) I605A mutant *Mt*DnaG-CTD. The mutation of a single residue, Ile605 to Ala, on the helicase-binding surface resulted in an almost 10-fold decrease in the binding affinity.

Fig. 6: *In silico* investigations of the cross-species helicase-primase complexes. (a-c) Experimentally determined structures of *Mt*DnaB-NTD (PDB: 2R5U) and (a) *Ec*DnaG-CTD (PDB: 1T3W), (b) *Hp*DnaG-CTD (PDB: 4EHS) and (c) *Vc*DnaG-CTD (PDB: 4IM9) were used to generate models of the helicase-primase complex. In each case, the crystal structure of the *Bst*DnaB-*Bst*DnaG-CTD (PDB: 2R6A) complex was used as a template. The model was further relaxed to eliminate bad contacts using molecular energy minimization with the help of the GROMACS molecular dynamics package. (d) RMSD analyses of the trajectories of the three models. Time-dependent Cα-RMSDs for all residues of DnaB-NTD complexed with *Ec*DnaG-CTD (red), *Hp*DnaG-CTD (black) and *Vc*DnaG-CTD (green) showed structural convergence from the initial time point.There were no structural changes during the dynamics simulation, but the DnaB-DnaG binding affinities for the cross-species complexes differed significantly from that for the native *M. tuberculosis* DnaB-DnaG complex.

Fig. 7: SPR sensorgrams for measuring the binding affinities of helicase with non-cognate primases. Sensorgrams showing the binding patterns of *Mt*DnaB-NTD at various concentrations (legends) with immobilized (a) *E. coli* DnaG-CTD, (b) *H. pylori* DnaG-CTD and (c) *V. cholerea* DnaG-CTD. The results indicated *Mt*DnaB-NTD to weakly bind *Ec*DnaG-CTD and *HpDn*aG-CTD, and unstably interact with *Vc*DnaG-CTD.

Figures:

Fig. 1

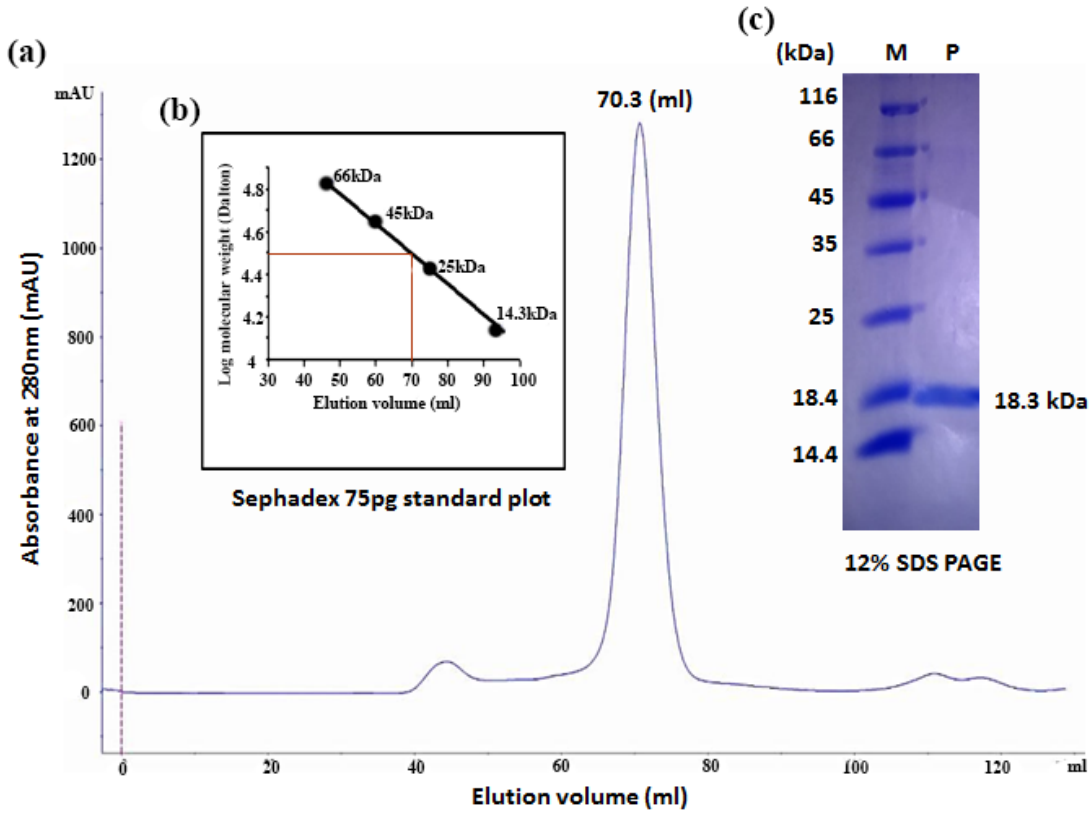


Fig. 2

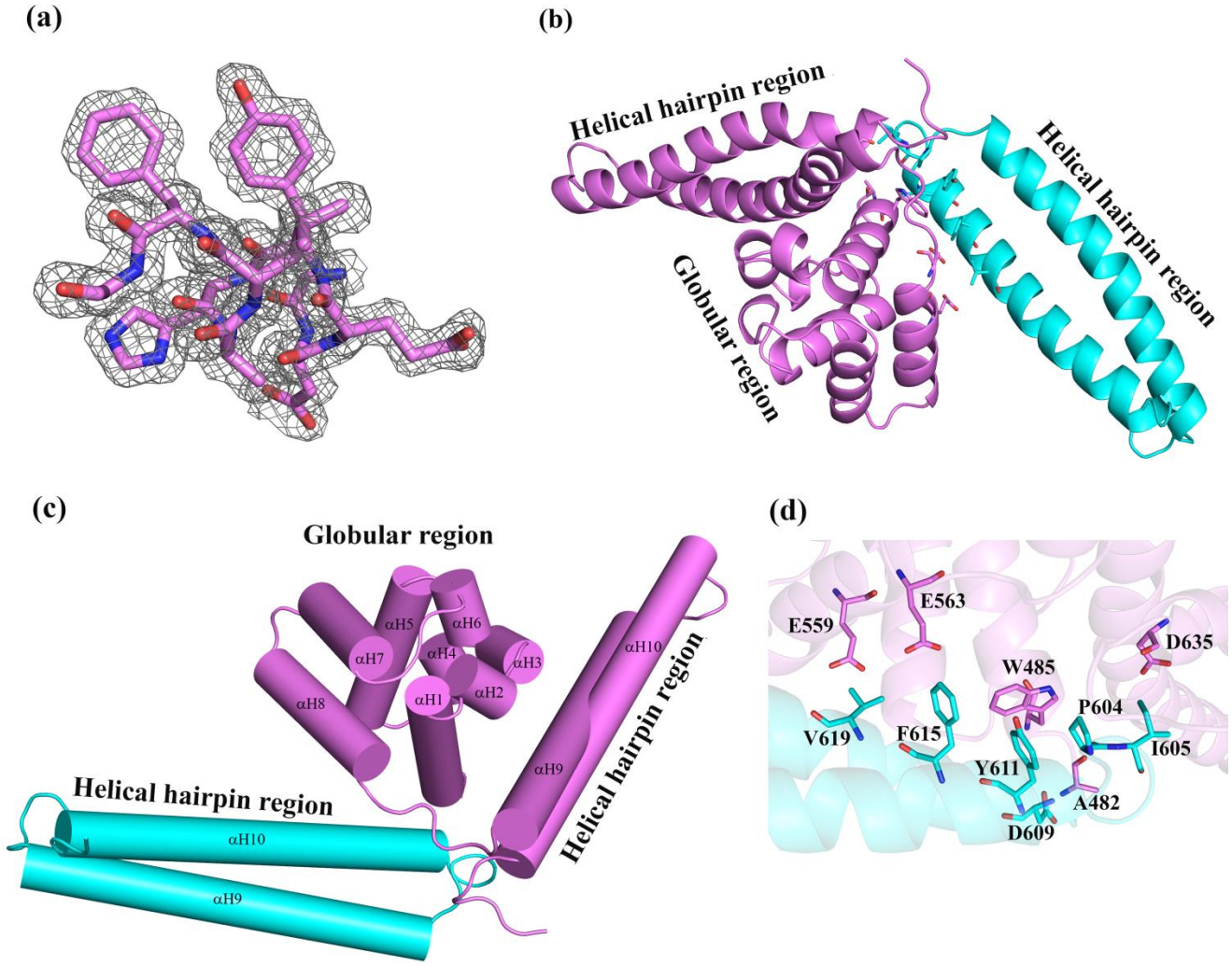

(a)
(b)
Helical hairpin region
Helical hairpin region
Globular region
(c)
Globular region
αH5
αH6
αH7
αH4
αH3
αH8
αH1
αH2
αH10
αH9
Helical hairpin region
Helical hairpin region
αH10
αH9
(d)
E559
E563
D635
W485
P604
V619
F615
I605
Y611
A482
D609

Fig. 3

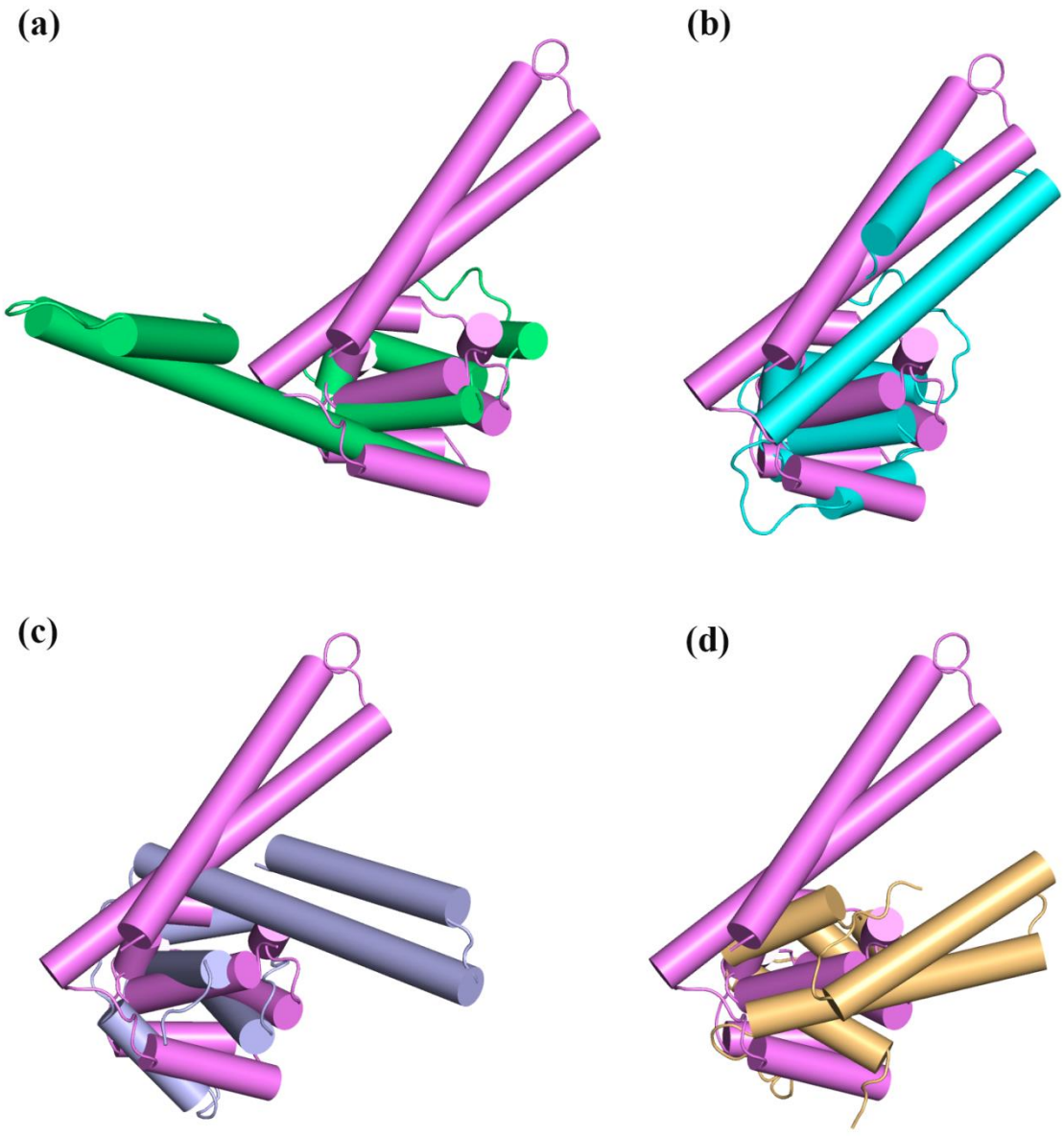


Fig. 4

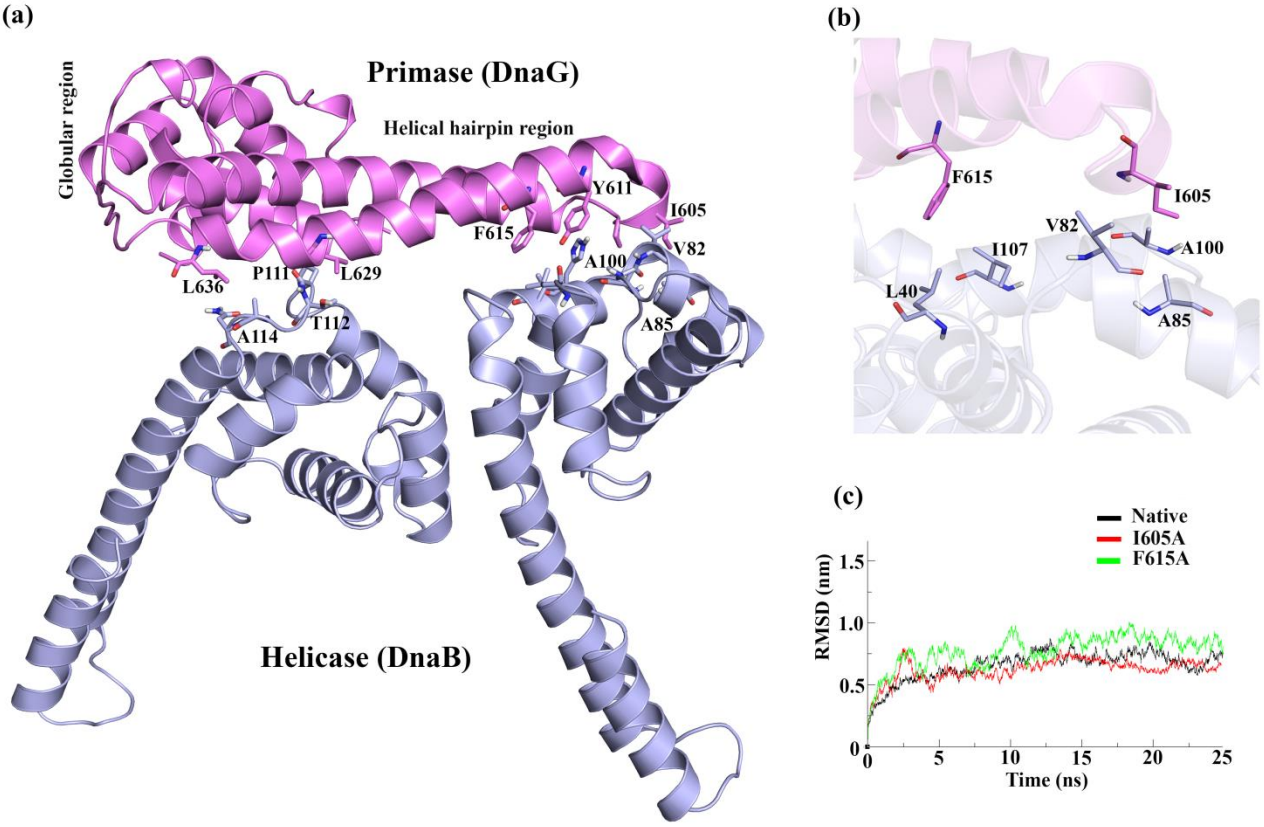

(a)
Primase (DnaG)
Globular region
Helical hairpin region
Y611
I605
F615
V82
A100
L636
P111
L629
A114
T112
A85
Helicase (DnaB)
(b)
F615
I605
V82
I107
A100
L40
A85
(c)
Native
I605A
F615A
RMSD (nm)
1.5
1.0
0.5
0
0
5
10
15
20
25
Time (ns)

Fig. 5

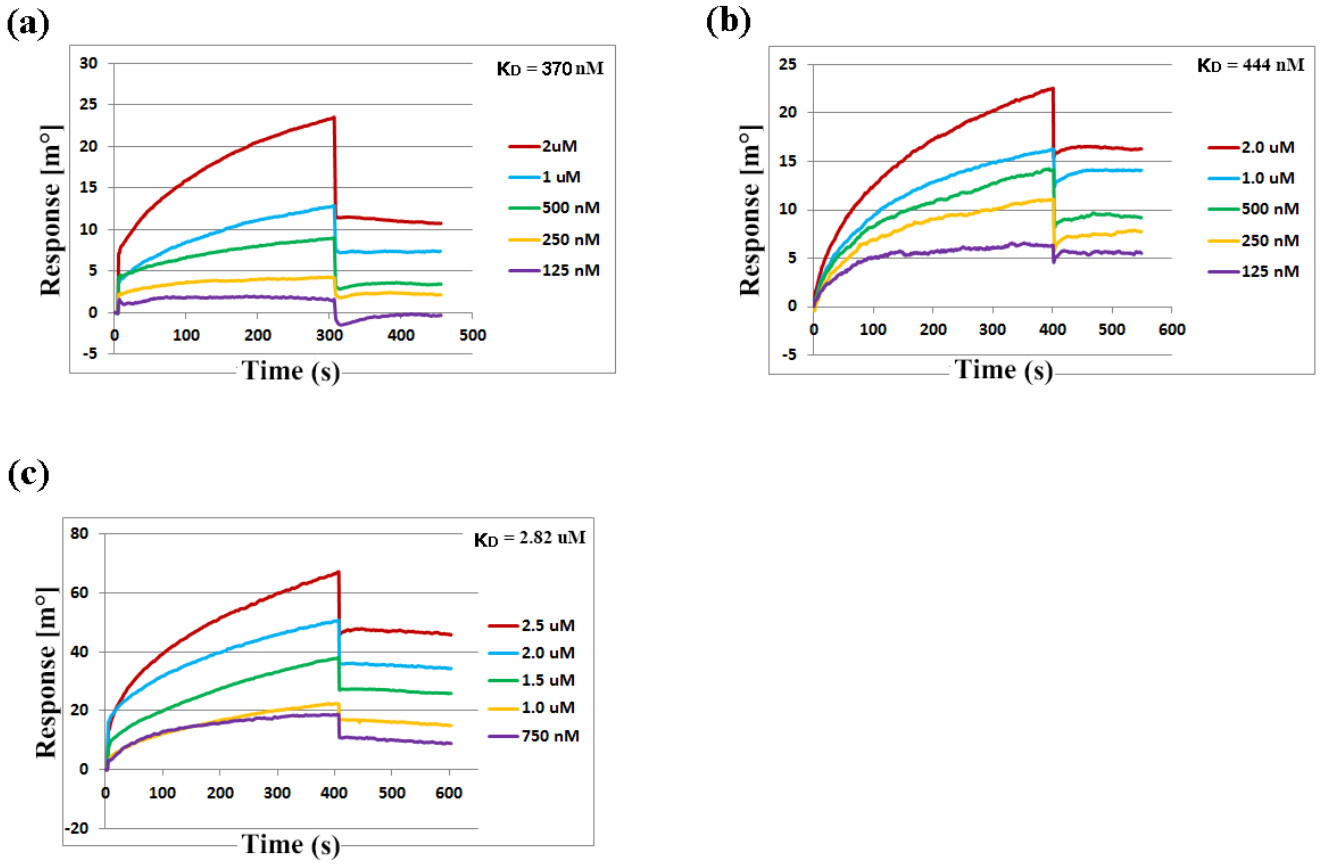
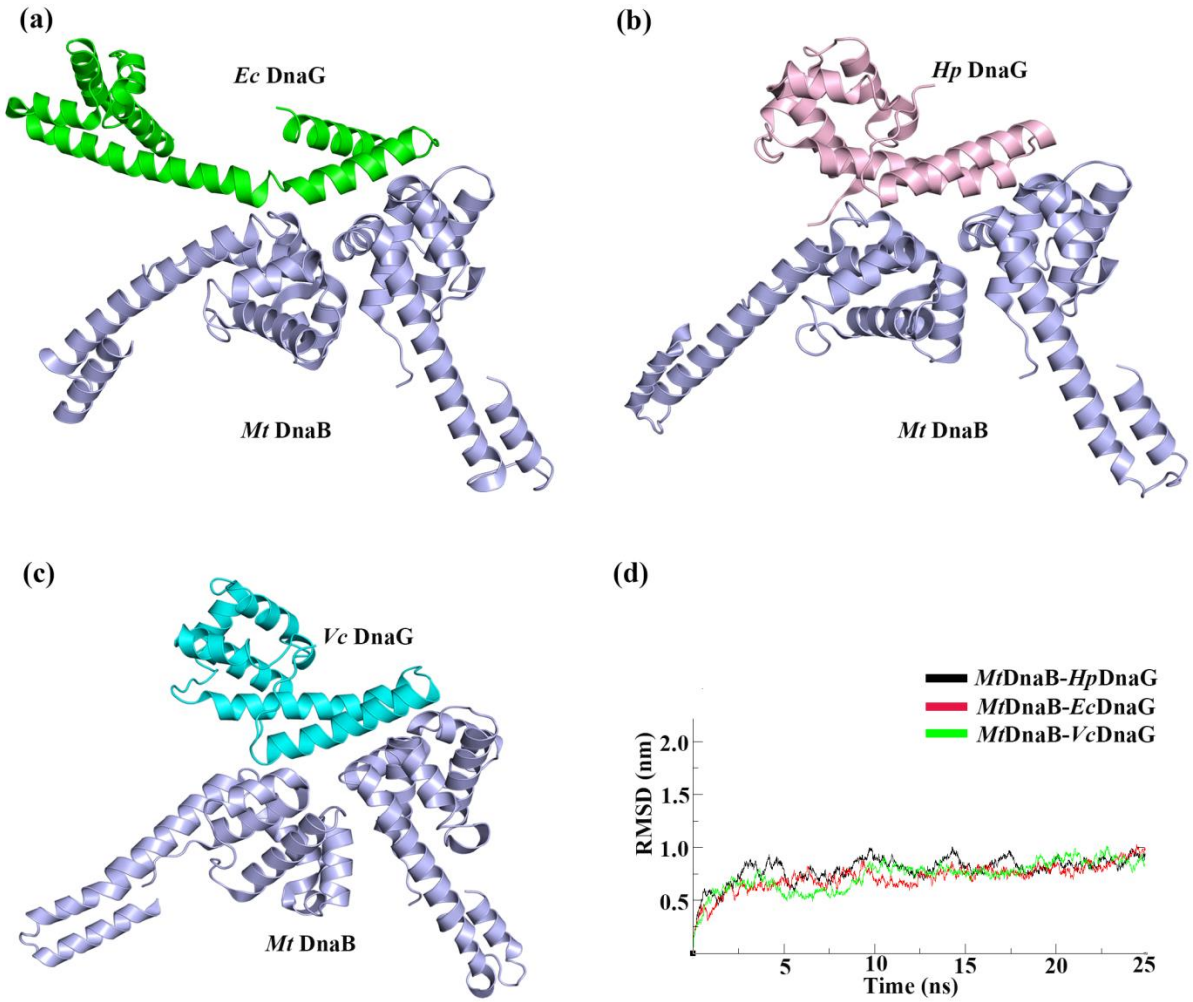

(a)
KD = 370 nM
2uM
1 uM
500 nM
250 nM
125 nM
Response [m°]
Time (s)
(b)
KD = 444 nM
2.0 uM
1.0 uM
500 nM
250 nM
125 nM
Response [m°]
Time (s)
(c)
KD = 2.82 uM
2.5 uM
2.0 uM
1.5 uM
1.0 uM
750 nM
Response [m°]
Time (s)

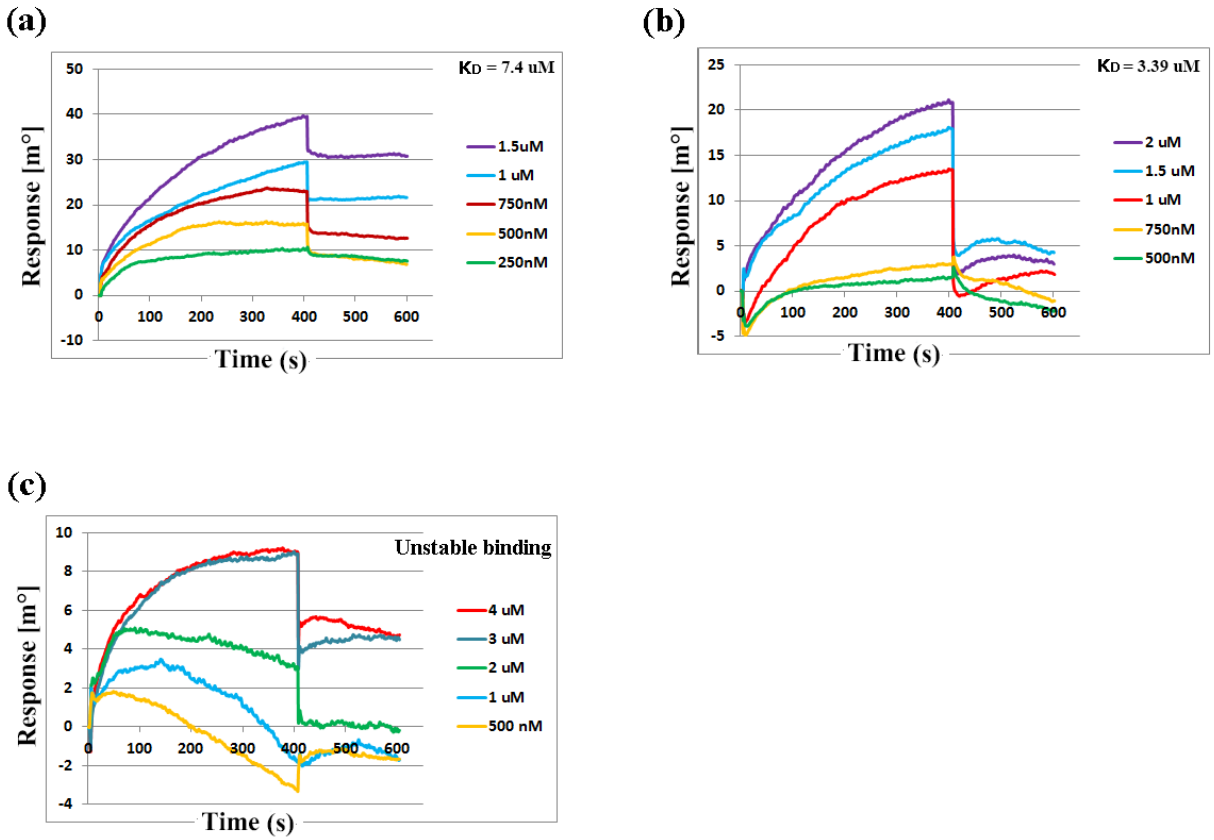
Fig. 6


(a)
Ec DnaG
Mt DnaB
(b)
Hp DnaG
Mt DnaB
(c)
Vc DnaG
Mt DnaB
(d)
MtDnaB-HpDnaG
MtDnaB-EcDnaG
MtDnaB-VcDnaG
RMSD (nm)
2.0
1.5
1.0
0.5
5
10
15
20
25
Time (ns)

906

907

908 **Fig. 7**



909

910

911

912

913

914

Supporting Data

Supplementary Fig. 1: DnaG primase. (a) Domain organization of DnaG primase. DnaG has three domains: zinc-binding domain for DNA binding, polymerase domain for primer synthesis, and C-terminal domain for helicase binding. The sequences of the zinc-binding domain and polymerase domain are conserved in DnaG primases, but the helicase-binding domain shows high variation. (b) Multiple sequence alignment of DnaG-CTD or helicase-binding domain shows insignificant homology.

Supplementary Fig. 2: Cleavage of *Mt*DnaG-CTD into two fragments. (a) About a week old purified *Mt*DnaG-CTD protein was loaded into 15% SDS-PAGE, yielded three bands, indicating part of the protein sample was cleaved into two fragments. This experiment confirms the crystal structure, where cleaved HHR fragment and DnaG-CTDs assembled into one assymmetric unit. The three bands indicate full-length CTD, lower twin bands represent globe region and HHR. (b) MALDI analysis of all three bands confirmed the cleavage of *Mt*DnaG-CTD into two fragments.

Supplementary Fig. 3: SPR sensorgram: (a) SPR sensorgram for measuring the binding pattern of BSA with different concentrations of *Mt*DnaB-NTD as a negative control. SPR sensorgrams showing the binding pattern of *Mt*DnaG-CTD with different concentrations of (b) monomeric *Mt*DnaB-NTD and (c) dimeric *Mt*DnaB-NTD. Only dimeric DnaG-CTD showed a pattern indicative of binding DnaB-NTD.

Supplementary Fig. 4: Cartoons showing the helices in crystal structures of the DnaG primase C-terminal domain. Cylindrical representation of (a) *Ec*DnaG-CTD, (b) *Bst*DnaG-CTD, (c) *Hp*DnaG-CTD and (d) *Vc*DnaG-CTD showing the total number of helices present in each of the different DnaG-CTDs.

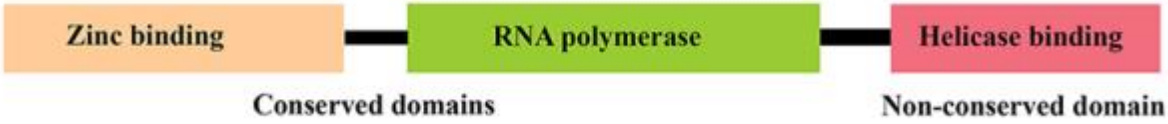
*X is very small (three residues) and it was not considered as a complete helix.

Supplementary Fig. 5: The helical hairpin region adopts different orientations in the free and bound states. Alignments of the globular regions (GRs) of the DnaG-CTDs gave insight into the differences in the shapes and orientations of the HHRs. (a) Superposition of the *Bst*DnaG-CTD portion of the crystal structure of the *Bst*DnaB-*Bst*DnaG-CTD complex (blue) with the NMR structure of (uncomplexed) *Bst*DnaG-CTD (green), showing a difference of about 65° between the orientations of their hairpins. (b) Superposition of the (uncomplexed) *Mt*DnaG-CTD crystal structure (light blue) with the *Mt*DnaG-CTD portion of the model of the complex (orange), showing a difference of about 32° between their hairpin orientations. (c) Structure of the *Bst*DnaB/*Bst*DnaG-CTD complex (PDB: 2R6A) used as a template for the *Mt*DnaB/*Mt*DnaG model. (d) Superposition of the structure of *Mt*DnaB-NTD on that of the *Bst*DnaB-NTD domain, and a superposition of the *Mt*DnaG-CTD GR structure on the *Bst*DnaG-CTD GR portion of the structure of the *Bst*DnaB-*Bst*DnaG-CTD complex. For superimposing the CTD helical hairpin structures, the helical hairpin of *Mt*DnaG-CTD was rotated by about 32°, keeping the linker between the HHR and the GR as the center of rotation, in order to obtain the lowest root mean square deviation (RMSD) between the *Bst*DnaG-CTD and *Mt*DnaG-CTD structures.

Supplementary Fig. 1

(a)

DnaG-primase domain organization



Conserved domains

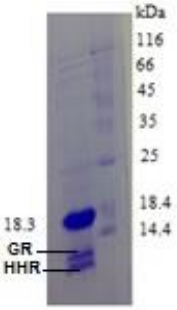
Non-conserved domain

(b)

```
M.tuberculosis          --------------------MRDPTLWPQREAL-----------KSALQYPALAGPVFD
B.stearothermophilus    --------------------MA-KKLLPAFQNAERLLLAHMMRSRDVALVVQERIGGRFN
H.pylori                ERVSFQPFYPKTEKPNRPQRFAHVSSAPSLEFLEKLVIRYLLEDRSLLDLAVGYIHS---
E.coli                  -----------AAESGVSRPVPQLKRTTM-----RILIGLLVQNPELATLVPPLENL---
V.cholerae              ---------------NRPAPHKAIKRTPM-----RDVIALLIQNPSYAELVPDLASV---
                                              .                   .

M.tuberculosis          ALTVEGFTHPEYAAVRAA---IDTA-GGTSAGLSGAQWLDMVRQQTTSTVTSALISELGV
B.stearothermophilus    IEEHRALAAYIYAFYEEG---HEADPGALISRIPGE--------------LQPLASELSL
H.pylori                -----GVFLHKKQEFDALCQEKLDDPK-----LVA-----------------LLLDANLP
E.coli                  ----DENKLPGLGLFRELVNTCLSQPG-----LTTGQLLEHYRGTNNAATLEKLSMWDDI
V.cholerae              ----RHLMIPGLDTFSEVLEKCRQYPH-----ITTGQLLEHWRDSKNETLLSRLASWEIP
                                                    :                   *

M.tuberculosis          --EAIQVDDDKLPRYIAGVLARLQEVWLGRQIAEVKSK-------LQRMSPIEQGDEYHA
B.stearothermophilus    LLIADDVSEQELEDYIRHVLNRP--KWLMLKVKEQEK------------TEAERRKDFLT
H.pylori                --LKK----GGFEKELRLLILRY----FERQLKEIPKSSLPFSEKMICLKKA-RQAIMKL
E.coli                  --ADKNIAEQTFTDSLNHMFDSL----LELRQEELI-----ARERTHGLSNEERLELWTL
V.cholerae              --LVEDNQEELFLDSLDKILAQC----VEKQIENLQ-----AKERSVGLSTEEKRELQDL
                                  :   :  ::       .  :  :                .   :

M.tuberculosis          ---LFGDLVAMEAYRRSLLEQASGDDLT
B.stearothermophilus    AARIAKEMIEMKKMLSS----S------
H.pylori                ---KQGELVAI-------LE--------
E.coli                  ---NQELA-KK-----------------
V.cholerae              ---ILKGL-KA-------LE--------
```

Supplementary Fig. 2

(a)



15% SDS-PAGE

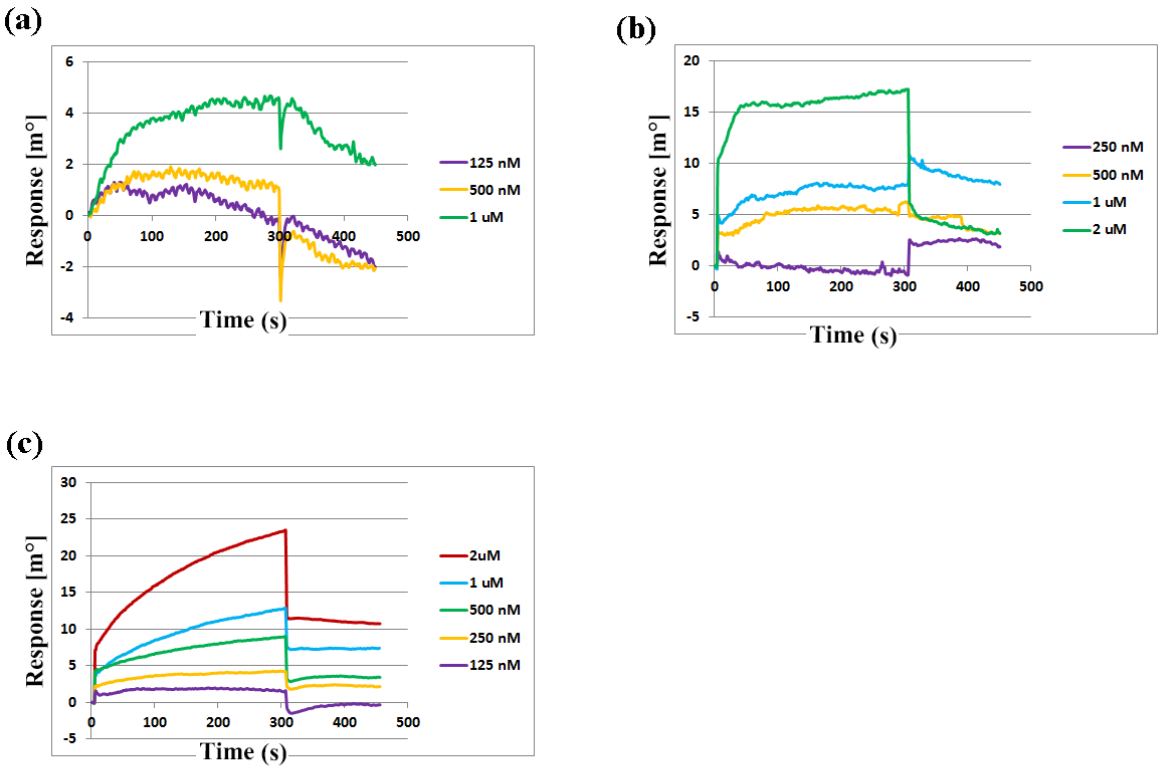
(b)

	Accession	Mass	Score	Description
1.	gi|289762510	16940	101	DNA primase dnaG [Mycobacterium tuberculosis GM 1503]
2.	gi|15609480	70006	60	DNA primase [Mycobacterium tuberculosis H37Rv]
3.	gi|326903959	69991	60	DNA primase dnaG [Mycobacterium tuberculosis W-148]
4.	gi|121638231	70105	48	DNA primase [Mycobacterium bovis BCG str. Pasteur 1173P2]
5.	gi|16588436	5537	42	RNA polymerase beta subunit [Mycobacterium tuberculosis]
6.	gi|294764463	6545	38	RNA polymerase beta subunit [Mycobacterium tuberculosis]
7.	gi|294764469	6503	38	RNA polymerase beta subunit [Mycobacterium tuberculosis]

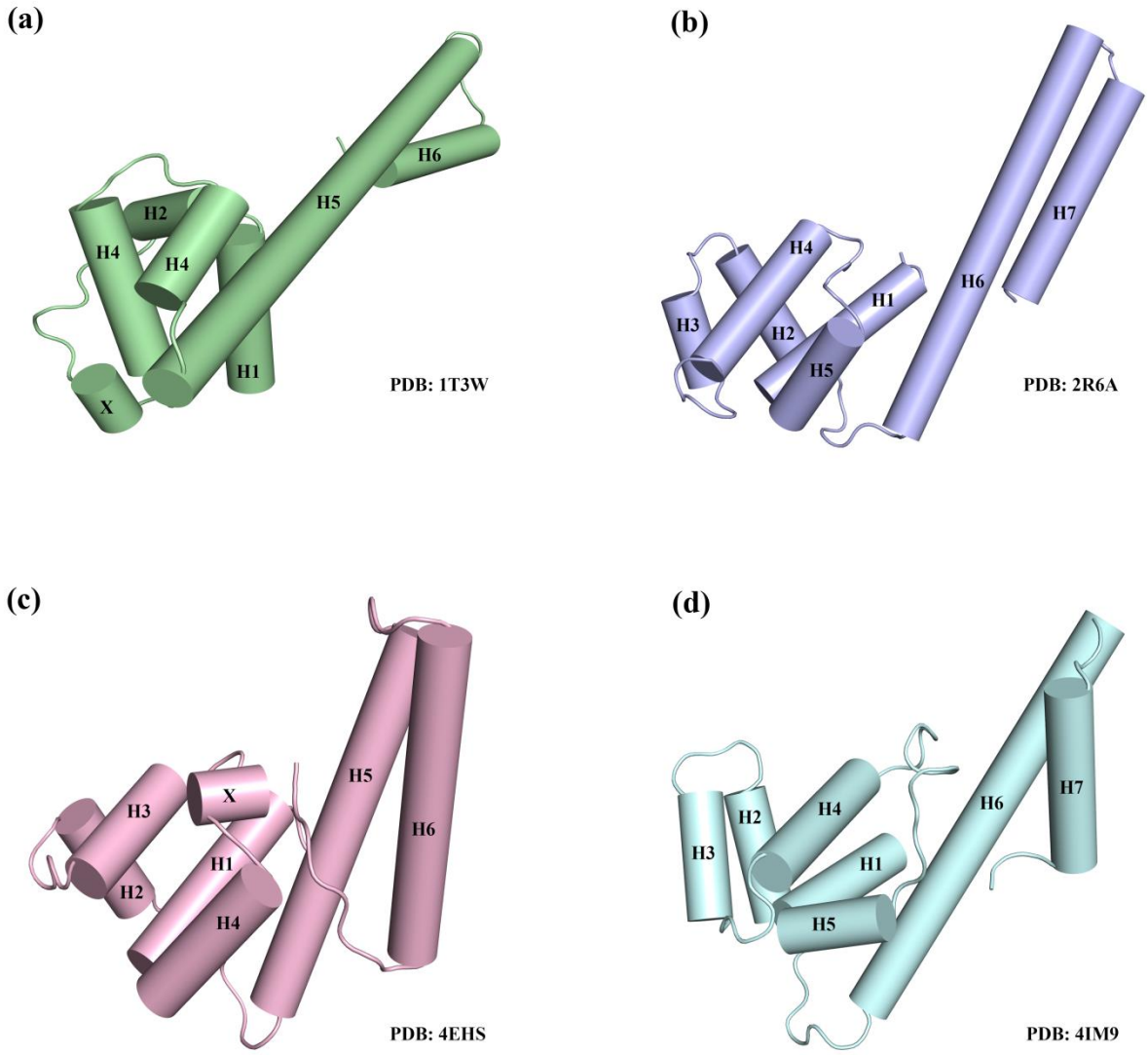
	Accession	Mass	Score	Description
1.	gi|289762510	16940	187	DNA primase dnaG [Mycobacterium tuberculosis GM 1503]
2.	gi|15609480	70006	160	DNA primase [Mycobacterium tuberculosis H37Rv]
3.	gi|326903959	69991	160	DNA primase dnaG [Mycobacterium tuberculosis W-148]
4.	gi|121638231	70105	92	DNA primase [Mycobacterium bovis BCG str. Pasteur 1173P2]
5.	gi|219555968	36451	49	MCE-family protein mce1B [Mycobacterium tuberculosis T17]
6.	gi|15607311	37818	49	MCE-family protein MCE1B [Mycobacterium tuberculosis H37Rv]
7.	gi|31791348	37806	49	MCE-family protein MCE1B [Mycobacterium bovis AF2122/97]

	Accession	Mass	Score	Description
1.	gi|289762510	16940	158	DNA primase dnaG [Mycobacterium tuberculosis GM 1503]
2.	gi|15609480	70006	146	DNA primase [Mycobacterium tuberculosis H37Rv]
3.	gi|326903959	69991	146	DNA primase dnaG [Mycobacterium tuberculosis W-148]
4.	gi|121638231	70105	82	DNA primase [Mycobacterium bovis BCG str. Pasteur 1173P2]
5.	gi|215447179	8895	53	signal recognition particle protein ffh [Mycobacterium tuberculosis T8
6.	gi|15610053	55024	44	signal recognition particle protein [Mycobacterium tuberculosis H37Rv]
7.	gi|253797996	55028	44	signal recognition particle protein ffh [Mycobacterium tuberculosis KZ

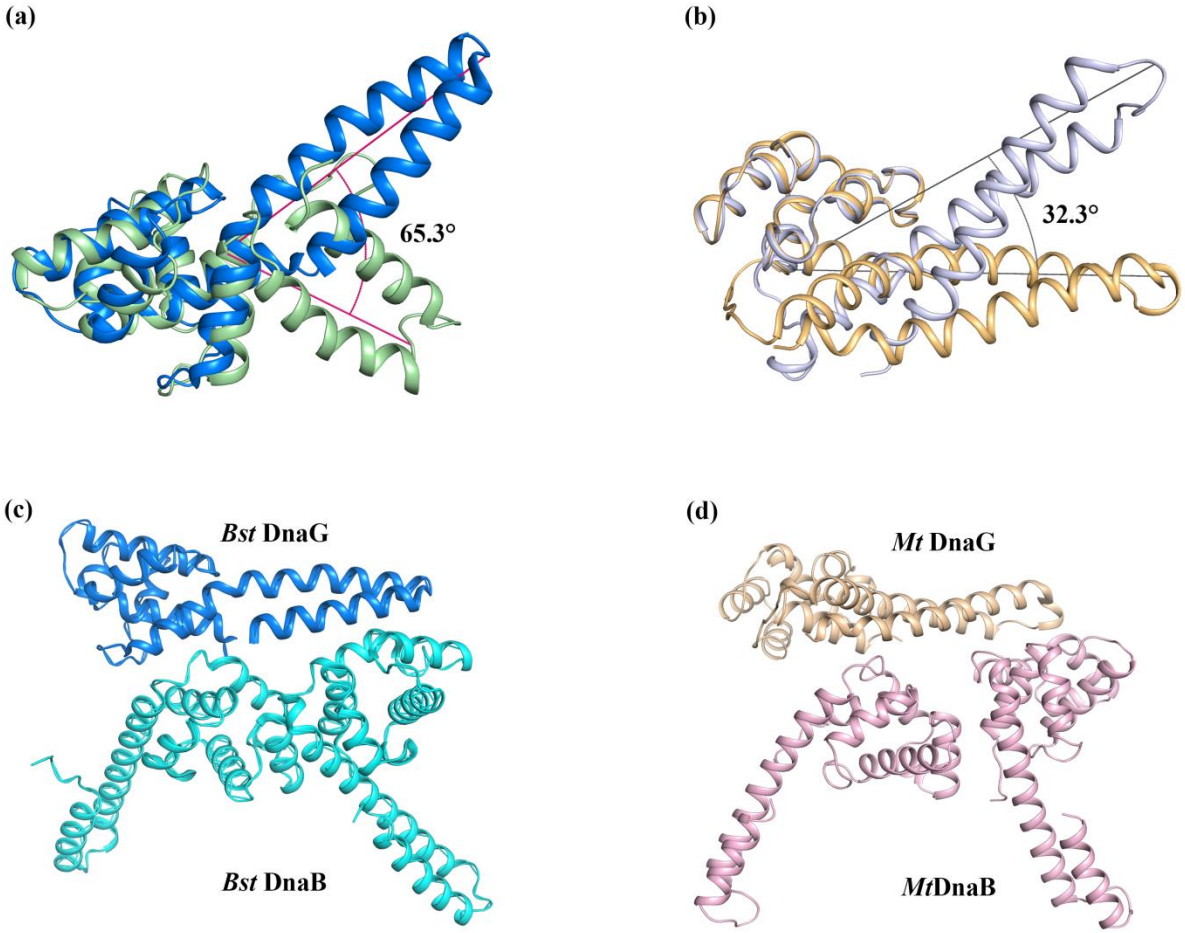
Supplementary Fig. 3



Supplementary Fig. 4



Supplementary Fig. 5



Supplementary Table 1: **List of primers used for the experiments.**

Fw-*Mt*DnaG-CTD	CTAGCTAGCACCCTGTGGCCACAGCGTGAGGCACTC
Rv-*Mt*DnaG-CTD	CGGCTCGAGGTCGCCGCTGGCCTGCTCTAACAGGCT
Fw-*Mt*DnaGI605A	CATGTCGCCGGCGGAGCAAGGCG
Rv-*Mt*DnaGI605A	CGCTGCAGTTTGGACTTC
Fw-*Mt*DnaGF615A	CCACGCGCTGGCGGGCGACCTGG
Rv-*Mt*DnaGF615A	TATTCGTCGCCTTGCTCG
Fw-*Mt*DnaB-NTD	GGAATTCCATATG ATG GGC CGT CAA CCA CCG CAG G
Rv *Mt*DnaB-NTD	CCGCTCGAG CTC TAC CAG CCG GCG CAG CA

*****Primers for DnaG-CTD of *H. pylori*, *E.coli*, and *V.cholerea* were listed previously.

Supplementary Table 2: Structure-based alignment of *M. tuberculosis* DnaG-CTD with other structures using Rapido

Name	#aligned residues	RMSD rigid	#aligned residues (rigid body)	RMSD flex	#rigid bodies
E. coli	101	12.58	69	0.52	5
B. stearothermophilus	131	6.82	87	1.01	4
H. pylori	108	5.24	85	0.90	4
V. cholerea	122	4.28	77	0.81	3

Supplementary Table 3: Average B factors of different regions of the DnaG C-terminal domain

Protein	**Globular region** ($Å^2$)	**Linker** ($Å^2$)	**Hairpin helix** ($Å^2$)
*Mt*DnaG (PDB)	17.89	25.18	19.28
*Ec*DnaG (IT3W)	61.70	89.60	66.10
*Bst*DnaG (2R6A)	98.92	103.06	92.64
*Hp*DnaG (4EHS)	18.4	24.30	22.70
*Vc*DnaG (4IM9)	33.87	51.67	39.74

Supplementary Table 4: Free binding energy values of the binding of native and mutant *M. tuberculosis* primases to native *M. tuberculosis* helicase.

No.	Complex	ΔE elec (kcal mol^{-1})	ΔE vdW (kcal mol^{-1})	ΔG polar (kcal mol^{-1})	ΔG nonpolar (kcal mol^{-1})	ΔG binding (kcal mol^{-1})
1.	Native *Mt*DnaG-CTD	-624.8±1.8	-92.3±1.3	960.2±4.0	-20.9±5.6	-100.3±3.7
2.	Mutant *Mt*DnaG-CTD (I605A)	-200.0±3.6	-87.8±1.7	906.9±3.8	-3.7±6.6	-60.9±2.4
3.	Mutant *Mt*DnaG-CTD (F615A)	-293.2±5.9	-91.9±1.5	913.8±6.82	-14.1±9.0	-75.3±3.2